Artificial Intelligence 175 (2011) 165–192

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Iterated belief change in the situation calculus ☆

Steven Shapiro [a,*], Maurice Pagnucco [b], Yves Lespérance [c], Hector J. Levesque [a]

[a] Department of Computer Science, University of Toronto, Toronto, ON M5S 3G4, Canada
[b] ARC Centre of Excellence for Autonomous Systems and National ICT Australia, School of Computer Science and Engineering, The University of New South Wales, Sydney, NSW 2052, Australia
[c] Department of Computer Science and Engineering, York University, Toronto, ON M3J 1P3, Canada

ARTICLE INFO

Article history:
Available online 3 April 2010

Keywords:
Knowledge representation and reasoning
Reasoning about action and change
Situation calculus
Belief change

ABSTRACT

John McCarthy's *situation calculus* has left an enduring mark on artificial intelligence research. This simple yet elegant formalism for modelling and reasoning about dynamic systems is still in common use more than forty years since it was first proposed. The ability to reason about action and change has long been considered a necessary component for any intelligent system. The situation calculus and its numerous extensions as well as the many competing proposals that it has inspired deal with this problem to some extent. In this paper, we offer a new approach to belief change associated with performing actions that addresses some of the shortcomings of these approaches. In particular, our approach is based on a well-developed theory of action in the situation calculus extended to deal with belief. Moreover, by augmenting this approach with a notion of plausibility over situations, our account handles nested belief, belief introspection, mistaken belief, and handles belief revision and belief update together with iterated belief change.

The work of John McCarthy has had a profound and lasting effect on artificial intelligence research. One of his more enduring contributions has been the introduction of the *situation calculus* [2,3]. This simple yet elegant formalism for modelling and reasoning about dynamic systems is still in common use more than forty years since it was first proposed. The ability to reason about action and change has long been considered a necessary component for any intelligent system. An agent acting in its environment must be capable of reasoning about the state of its environment and keeping track of any changes to the environment as actions are performed. Various theories have been developed to give an account of how this can be achieved. Foremost among these are theories of belief change and theories for reasoning about action. While originating from different motivations, the two are united in their aim to have agents maintain a model of the environment that matches the actual environment as closely as possible given the available information. An important consideration is the ability to deal with a succession of changes; known as the problem of *iterated belief change*.

In this paper, we consider a new approach for modelling iterated belief change using the language of the situation calculus [2,3]. While our approach is in some ways limited in its applicability, we feel that it is conceptually very simple and offers a number of useful features not found in other approaches:

- It is completely integrated with a well-developed theory of action in the situation calculus [4] and its extension to handle knowledge expansion [5,6]. Specifically, the manner in which beliefs change in our account is simply a special

☆ An earlier version of this paper appeared in Shapiro et al. (2000) [1].

* Corresponding author.
E-mail addresses: steven@cs.toronto.edu (S. Shapiro), morri@cse.unsw.edu.au (M. Pagnucco), lesperan@cse.yorku.ca (Y. Lespérance), hector@ai.toronto.edu (H.J. Levesque).

0004-3702/$ – see front matter
doi:10.1016/j.artint.2010.04.003

case of how other fluents change as the result of actions, and thus among other things, we inherit a solution to the frame problem.
- Like Scherl and Levesque [5,6], our theory accommodates both belief *update* and belief *expansion*. The former concerns beliefs that change as the result of the realization that the world has changed; the latter concerns beliefs that change as the result of newly acquired information.
- Unlike Scherl and Levesque, however, our theory is not limited to belief expansion; rather it deals with the more general case of belief *revision*. It will be possible in our model for an agent to believe some formula ϕ, acquire information that causes it to change its mind and believe $\neg\phi$ (without believing the world has changed), and later go back to believing ϕ again. In Scherl and Levesque and in other approaches based on this work such as [7,8], new information that contradicts previous beliefs cannot be consistently accommodated.
- Because belief change in our model is always the result of action, our account naturally supports *iterated* belief change. This is simply the result of a sequence of actions. Moreover, each individual action can potentially cause both an update (by changing the world) and a revision (by providing sensing information) in a seamless way.
- Like Scherl and Levesque and unlike many previous approaches to belief change, e.g., [9,10], our approach supports belief *introspection*: an agent will know what it believes and does not believe. Furthermore, it has information about the past, and so will also know what it used to believe and not believe. Finally, an agent will be able to predict what it will believe in the future after it acquires information through sensing.
- Unlike Scherl and Levesque, our agents will be able to introspectively tell the difference between an update and a revision as they move from believing ϕ to believing $\neg\phi$. In the former case, the agent will believe that it believed ϕ in the past, and that it was correct to do so; in the latter case, it will believe that it believed ϕ in the past but that it was *mistaken*.
- One important lesson learned is that not only does our method for iterated belief change in the situation calculus possess interesting properties but attempting to use more sophisticated schemes that involve modifying plausibilities of possible worlds, leads to unintuitive introspection properties when applied to situations.

The rest of the paper is organized as follows: in the next section, we briefly review the situation calculus including the Scherl and Levesque [5,6] model of belief expansion, and we review the most popular accounts of belief revision, belief update, and iterated belief change; in Section 3, we motivate and define a new belief operator as a modification to the one used by Scherl and Levesque; in Section 4, we prove some properties of this operator, justifying the points made above; in Section 5, we show the operator in action on a simple example, and how an agent can change its mind repeatedly; in Section 6, we analyze the extent to which our framework satisfies revision, update, and iterated revision postulates; in Section 7, we compare our framework to some of the existing approaches to belief change; and in the final section, we draw some conclusions and discuss future work.

1. Background

The basis of our framework for belief change is an action theory [4] based on the situation calculus [2,3], and extended to include a belief operator [5,6]. In this section, we begin with a brief overview of the situation calculus and follow it with a short review of belief change in sufficient detail to understand the contributions made in this paper.

1.1. The situation calculus

The situation calculus is a predicate calculus language for representing dynamically changing domains. A situation represents a snapshot of the domain. There is a set of initial situations corresponding to the ways the agent[1] believes the domain might be initially. The actual initial state of the domain is represented by the distinguished initial situation constant, S_0, which may or may not be among the set of initial situations believed possible by the agent. The term $do(a,s)$ denotes the unique situation that results from the agent performing action a in situation s. Thus, the situations can be structured into a set of trees, where the root of each tree is an initial situation and the arcs are actions.

Predicates and functions whose value may change from situation to situation (and whose last argument is a situation) are called *fluents*. For instance, we use the fluent $\textsc{InR}_1(s)$ to represent that the agent is in room R_1 in situation s. The effects of actions on fluents are defined using successor state axioms [4], which provide a succinct representation for both effect axioms and frame axioms [2,3]. For example, assume that there are only two rooms, R_1 and R_2, and that the action $\textsc{leave}$ takes the agent from the current room to the other room. Then, the successor state axiom for $\textsc{InR}_1$ is[2]:

$$\textsc{InR}_1\big(do(a,s)\big) \equiv \big(\big(\neg\textsc{InR}_1(s) \wedge a = \textsc{leave}\big) \vee \big(\textsc{InR}_1(s) \wedge a \neq \textsc{leave}\big)\big).$$

This axiom asserts that the agent will be in R_1 after doing some action if and only if either the agent is in R_2 ($\neg\textsc{InR}_1(s)$) and leaves it or the agent is currently in R_1 and the action is anything other than leaving it.

[1] The situation calculus can accommodate multiple agents, but for the purposes of this paper we assume that there is a single agent, and all actions are performed by that agent.

[2] We adopt the convention that unbound variables are universally quantified in the widest scope.

Moore [11] defined a possible-worlds semantics for a modal logic of knowledge in the situation calculus by treating situations as possible worlds. Scherl and Levesque [5,6] adapted the semantics to the action theories of Reiter [4]. The idea is to have an accessibility relation on situations, $B(s', s)$, which holds if in situation s, the situation s' is considered possible by the agent. Note that the order of the arguments is reversed from the usual convention in modal logic.

Levesque [8] introduced a predicate, $SF(a, s)$, to describe the result of performing the binary-valued sensing action a. $SF(a, s)$ holds if and only if the sensor associated with a returns the sensing value 1 in situation s. Each sensing action senses some property of the domain. The property sensed by an action is associated with the action using a *guarded sensed fluent axiom* [12]. For example, suppose that there are lights in R_1 and R_2 and that $\textsc{Light}_1(s)$ ($\textsc{Light}_2(s)$, respectively) holds if the light in R_1 (R_2, respectively) is on. Then:

$$\textsc{InR}_1(s) \supset \big(SF(\textsc{senseLight}, s) \equiv \textsc{Light}_1(s)\big)$$

$$\neg\textsc{InR}_1(s) \supset \big(SF(\textsc{senseLight}, s) \equiv \textsc{Light}_2(s)\big)$$

can be used to specify that the $\textsc{senseLight}$ action senses whether the light is on in the room where the agent is currently located.

Scherl and Levesque [5,6] defined a successor state axiom for B that shows how actions, including sensing actions, affect the beliefs of the agent. We use the same axiom (with some notational variation) here:

Axiom 1 *(Successor state axiom for B).*

$$B\big(s'', do(a, s)\big) \equiv \exists s'\big[B\big(s', s\big) \wedge s'' = do\big(a, s'\big) \wedge \big(SF\big(a, s'\big) \equiv SF(a, s)\big)\big].$$

The situations s'' that are B-related to $do(a, s)$ are the ones that result from doing action a in a situation s', such that the sensor associated with action a has the same value in s' as it does in s. This axiom is further illustrated in Fig. 2 and the explanatory text that follows. We will see in Section 2.2 how a belief operator can be defined in terms of this fluent.

There are various ways of axiomatizing dynamic applications in the situation calculus. Here we adopt a simple form of the guarded action theories described by De Giacomo and Levesque [12] consisting of: (1) successor state axioms[3] for each fluent (including B and pl introduced below), and guarded sensed fluent axioms for each action, as discussed above; (2) unique names axioms for the actions, and domain-independent foundational axioms (given below); and (3) initial state axioms, which describe the initial state of the domain and the initial beliefs of the agent.[4] For simplicity, we assume here that all actions are always executable and omit the action precondition axioms and references to a *Poss* predicate that are normally included in situation calculus action theories. These do not add any significant complexity to our approach but omitting them here allows us to focus on the key elements of our framework.

In what follows, we will use Σ to refer to a guarded action theory of this form. By a *domain-dependent fluent*, we mean a fluent other than B or pl, and a *domain-dependent formula* is one that only mentions domain-dependent fluents.

As part of every guarded action theory, we have unique names axioms for actions and foundational axioms. The unique names axioms for the actions state that distinct action function symbols correspond to different action functions. For every pair of distinct action functions, $\mathbf{a}_1$ and $\mathbf{a}_2$, we need an axiom of the following form:

Axiom 2.

$$\mathbf{a}_1(\vec{x}) \neq \mathbf{a}_2(\vec{y}).$$

Also, for an action function, $\mathbf{a}$, we need an axiom of the following form:

Axiom 3.

$$\mathbf{a}(\vec{x}) = \mathbf{a}(\vec{y}) \supset \vec{x} = \vec{y}.$$

This means that an action function applied to distinct arguments is mapped to different actions, i.e., all action functions are injective. If we have n action functions, we need $O(n^2)$ unique names axioms [13]. However, it would not be difficult to have them automatically generated from a list of the action names and arities.

We want the situations to be the smallest set generated by sequences of actions starting in an initial situation. We axiomatize the structure of the situations with *foundational axioms* based on the ones listed in Levesque et al. [14] and Pirri

[3] We could use the more general *guarded successor state axioms* of De Giacomo and Levesque [12], but regular successor state axioms suffice for the simple domain we consider here and for illustrating our approach.

[4] These are axioms that only describe initial situations. Reiter [4] has adopted S_0 as the only initial situation, but to formalize belief, we need additional initial situations representing the alternative scenarios consistent with the agent's initial beliefs.

and Reiter [15] for the language of the "epistemic situation calculus". We first define the initial situations to be those that have no predecessors:

$$Init(s') \stackrel{\text{def}}{=} \neg\exists a, s.s' = do(a, s).$$

We declare S_0 to be an initial situation.

Axiom 4.

$$Init(S_0).$$

We also need an axiom stating that *do* is injective.

Axiom 5.

$$do(a_1, s_1) = do(a_2, s_2) \supset (a_1 = a_2 \wedge s_1 = s_2).$$

The induction axiom for situations says that if a property P holds of all initial situations, and P holds for all successors of situation s if it holds for s, then P holds for all situations.

Axiom 6.

$$\forall P.\big[\big(\forall s.Init(s) \supset P(s)\big) \wedge \big(\forall a, s.P(s) \supset P\big(do(a, s)\big)\big)\big] \supset \forall s P(s).$$

We now define precedence for situations. We say that s *strictly precedes* s' if and only if there is a (non-empty) sequence of actions that take s to s'.

Axiom 7.

$$\forall s_1, s_2.s_1 \prec s_2 \equiv \big(\exists a, s.s_2 = do(a, s) \wedge (s_1 \preceq s)\big),$$

where $s_1 \preceq s_2 \stackrel{\text{def}}{=} s_1 = s_2 \vee s_1 \prec s_2$ denotes that s_1 *precedes* s_2.

1.2. Belief change

Before formally defining a belief operator in this language, we briefly review the notion of belief change as it exists in the literature. Belief change, simply put, aims to study the manner in which an agent's epistemic (belief) state should change when the agent is confronted by new information. In the literature,[5] there is often a clear distinction between two forms of belief change: *revision* and *update*. Both forms can be characterized by an axiomatic approach (in terms of rationality postulates) or through various constructions (e.g., epistemic entrenchment, possible worlds, etc.). The AGM theory [9] is the prototypical example of belief revision while the KM framework [10] is often identified with belief update.

Intuitively speaking, belief revision is appropriate for modelling static environments about which the agent has only partial and possibly incorrect information. New information is used to fill in gaps and correct errors, but the environment itself does not undergo change. Belief update, on the other hand, is intended for situations in which the environment itself is changing due to the performing of actions.

For completeness and later comparison, we list here the AGM postulates [16,9] for belief revision. By $K * \phi$ we mean the revision of belief state K by new information ϕ.[6]

(K*1) $K * \phi$ is deductively closed
(K*2) $\phi \in K * \phi$
(K*3) $K * \phi \subseteq K + \phi$
(K*4) If $\neg\phi \notin K$, then $K + \phi \subseteq K * \phi$
(K*5) $K * \phi = \mathcal{L}$ iff $\models \neg\phi$
(K*6) If $\models \phi \equiv \psi$, then $K * \phi = K * \psi$
(K*7) $K * (\phi \wedge \psi) \subseteq (K * \phi) + \psi$
(K*8) If $\neg\psi \notin K * \phi$, then $(K * \phi) + \psi \subseteq K * (\phi \wedge \psi)$

---

[5] We shall restrict our attention to approaches in the AGM vein [16,9,10] although there are many others.

[6] In the AGM theory, K is a set of formulae and ϕ is a formula taken from an object language $\mathcal{L}$ containing the standard Boolean connectives and the logical constant $\perp$ (falsum). Furthermore, K is a set of formulae (from $\mathcal{L}$) closed under the deductive consequence operator Cn associated with the underlying logic. The operation $K + \phi$ denotes the belief expansion of K by ϕ and is defined as $K + \phi = Cn(K \cup \{\phi\})$. $[K]$ denotes the set of all consistent complete theories of $\mathcal{L}$ containing K.

Katsuno and Mendelzon [10] provide the following postulates for belief update, where $K \diamond \phi$ denotes the update of belief state K by formula ϕ.[7]

(K $\diamond$ 1) $K \diamond \phi$ is deductively closed
(K $\diamond$ 2) $\phi \in K \diamond \phi$
(K $\diamond$ 3) If $\phi \in K$, then $K \diamond \phi = K$
(K $\diamond$ 4) $K \diamond \phi = \mathcal{L}$ iff $K \models \bot$ or $\phi \models \bot$
(K $\diamond$ 5) If $\models \phi \equiv \psi$, then $K \diamond \phi = K \diamond \psi$
(K $\diamond$ 6) $K \diamond (\phi \wedge \psi) \subseteq (K \diamond \phi) + \psi$
(K $\diamond$ 7) If K is complete and $\neg\psi \notin K \diamond \phi$, then $(K \diamond \phi) + \psi \subseteq K \diamond (\phi \wedge \psi)$
(K $\diamond$ 8) If $[K] \neq \emptyset$, then $K \diamond \phi = \bigcap_{w \in [K]} w \diamond \phi$

One of the major issues in this area is that of *iterated belief change*, i.e., modelling how the agent's beliefs change after a succession of belief revisions or updates occur. Two of the main developments in this area are the work of Darwiche and Pearl [18] and Boutilier [19]. Darwiche and Pearl put forward the following postulates as a way of extending the AGM revision postulates to handle *iterated revision*.[8]

(DP1) If $\psi \models \phi$, then $(K * \phi) * \psi = K * \psi$
(DP2) If $\psi \models \neg\phi$, then $(K * \phi) * \psi = K * \psi$
(DP3) If $\phi \in K * \psi$, then $\phi \in (K * \phi) * \psi$
(DP4) If $\neg\phi \notin K * \psi$, then $\neg\phi \notin (K * \phi) * \psi$

In Section 5, we return to consider the extent to which our framework satisfies these postulates.

2. Our account of belief change

2.1. Belief change and introspection

Scherl and Levesque provide an elegant framework for incorporating knowledge change into the situation calculus. However, in many applications, there is a need to represent information that could turn out to be wrong, i.e., we need to be able to represent beliefs and how they change due to actions. In order to incorporate belief change into our framework, we decided to adapt ideas from Spohn [20] and Darwiche and Pearl [18]. Our first attempt was to add an extra argument to the accessibility relation. This extra argument was a natural number corresponding to the plausibility of the accessible situation.[9] $B(s', n, s)$ would denote that in s, the agent thinks s' was possible with κ-ranking (plausibility) n.[10] The lower κ-ranking, the more plausible the situation would be considered by the agent, and the beliefs of the agent in s would be determined by the situations accessible from s with κ-ranking 0, i.e.,

$$Bel(\phi, s) \stackrel{\text{def}}{=} \forall s'.B\big(s', 0, s\big) \supset \phi\big[s'\big].$$

The successor state axiom for B would adjust the plausibilities of the B-related situations depending on the results of sensing using a scheme similar to Darwiche and Pearl's. Unlike most other approaches to belief revision, we wanted to handle positive and negative introspection of beliefs. However, we realized that this desideratum was in conflict with any reasonable scheme for updating plausibilities of accessible situations. Any reasonable scheme would have the following property: *if the accessible situation agrees with the actual situation on the result of a sensing action, the plausibility of the situation should increase* (*i.e., its κ-ranking should decrease*), *otherwise the plausibility should decrease*. In other words, if $B(s', n, s)$ and $SF(a, s') \equiv SF(a, s)$ hold, then $B(do(a, s'), m, do(a, s))$ should hold for some $m \leqslant n$. Similarly, if $SF(a, s') \equiv \neg SF(a, s)$ holds, then m should be greater than or equal to n. On the other hand, to ensure positive and negative introspection of beliefs, we combined and generalized the constraints that B be transitive and Euclidean to obtain the following requirement on B (which we call (TE) for transitive and Euclidean):

$$\forall s, s'.\big(\exists n.B\big(s', n, s\big)\big) \supset \big[\forall s'', m.B\big(s'', m, s'\big) \equiv B\big(s'', m, s\big)\big]. \tag{TE}$$

This requirement ensures that any situation s' accessible from s has the same belief structure as s, i.e., s' has the same accessible situations with the same plausibilities as s. This ensures that the agent has positive and negative introspection of

[7] To facilitate comparison with the AGM postulates, we have reformulated the original postulates of Katsuno and Mendelzon into an equivalent set using AGM-style terminology [17]. For renderings of these postulates and the AGM postulates above in the KM-style, refer to Katsuno and Mendelzon [10].

[8] Again, for consistency of presentation, we have translated the Darwiche and Pearl postulates into AGM-style terminology rather than KM-style terminology used in the original paper.

[9] In fact, the actual numbers assigned to the situations are not relevant. All that is important is the ordering of the situations by plausibility. We could have used any total preorder on situations for this purpose, but using $\leqslant$ on natural numbers simplifies the presentation of our framework.

[10] We adopt Spohn's [20] terminology here.

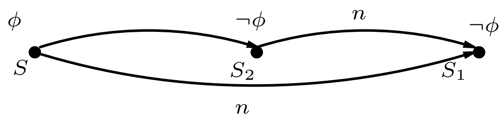

Fig. 1. Belief introspection and plausibility update clash.

its entire epistemic state, i.e., both its beliefs and conditional beliefs. For example, in Fig. 3, all the situations within each oval are mutually accessible (and only these situations are accessible) with the plausibility indicated inside the oval.

To see why this requirement conflicts with any reasonable plausibility update scheme for perfectly accurate sensors, consider the following example illustrated in Fig. 1. In the figure, there are three situations, S, S_1, and S_2. S_1 and S_2 are accessible from S, and S_1 has κ-ranking n (the κ-ranking of S_2 is irrelevant to the example). In order to satisfy (TE), S_1 must also be accessible from S_2 with κ-ranking n. However, after the agent senses ϕ (i.e., after it performs the action $sense_\phi$), S_1 will have to become less plausible relative to S, but more plausible relative to S_2, since S_1 and S disagree on the value of ϕ, whereas S_1 and S_2 agree. Therefore, (TE) will in general be violated after the agent senses ϕ.

A possible solution to this problem is to consider only what the agent believes in the actual situation, i.e., in S_0 and its successors, and set the plausibilities of all accessible situations according to whether they agree with the actual situation on the value of the property being sensed. In our example, if we take S to be S_0, S_1 would then become less plausible relative to *both* S and S_2. Unfortunately, this solution can also lead to subtle undesirable introspective properties. For instance, in Section 6 we show how one can construct an example where the following holds: $Bel(\neg\phi \wedge Bel(\phi, do(sense_\phi, now)), S)$, i.e., in S, the agent believes $\neg\phi$, and also believes that after sensing ϕ, it will believe ϕ, which is counterintuitive. If the agent believes $\neg\phi$, then it should also believe that it will continue to believe $\neg\phi$ after sensing ϕ.

Our resolution to this problem, which is discussed at length in the following sections, was to revert back to a binary accessibility relation and to use Scherl and Levesque's successor state axiom for B. Instead of assigning plausibilities relative to a situation, each situation is assigned an absolute plausibility using a functional fluent $pl(s)$, which maps a situation to a natural number corresponding to the κ-ranking of s (again, the lower the κ-ranking, the more plausible the situation). The plausibilities of successor situations are constrained to be the same as their predecessors, i.e., the plausibility of a situation is unaffected by actions. The beliefs of the agent in a situation s are those formulae true in the most plausible situations accessible from s, but these situations are no longer required to have κ-rank 0. When sensing occurs, accessible situations will be dropped, therefore the set of the most plausible accessible situations will change, and the agent's beliefs will change. Since Scherl and Levesque's successor state axiom for B preserves (TE), positive and negative introspection will be maintained, if it holds initially.

2.2. Definition of the belief operator

In this section, we define what it means for an agent to believe a formula ϕ in a situation s, i.e., $Bel(\phi, s)$. Since ϕ will usually contain fluents, we introduce a special symbol *now* as a placeholder for the situation argument of these fluents, e.g., $Bel(\textsc{InR}_1(now), s)$. $\phi[s]$ denotes the formula that results from substituting s for *now* in ϕ. To make the formulae easier to read, we will often suppress the situation argument of fluents in the scope of a belief operator, e.g., $Bel(\textsc{InR}_1, s)$.

Scherl and Levesque [5,6], define a modal operator for belief in terms of the accessibility relation on situations, $B(s', s)$. For Scherl and Levesque, the believed formulae are the ones true in all accessible situations:

Definition 8.

$$Bel_{\mathrm{SL}}(\phi, s) \overset{\mathrm{def}}{=} \forall s'\big(B\big(s', s\big) \supset \phi\big[s'\big]\big).$$

To understand how belief change works, both in Scherl and Levesque and here, consider the example illustrated in Fig. 2. In this example, we have three initial situations S, S_1, and S_2 (across the bottom of the diagram). S_1 and S_2 are B-related to S (i.e., $B(S_1, S)$ and $B(S_2, S)$), as indicated by the arrows labelled B. (Ignore the circles around certain situations for now.) In all three situations, the agent is not in the room R_1. In S and S_2 the light in R_1 is on, and in S_1 the light is off. So at S, the agent believes it is not in R_1 (i.e., that it is in R_2), but it has no beliefs about the status of the light in R_1. We first consider the action of leaving R_2, which will lead to a belief update. By the successor state axiom for B, both $do(\textsc{leave}, S_1)$ and $do(\textsc{leave}, S_2)$ are B-related to $do(\textsc{leave}, S)$. In the figure, these three situations are called S'_1, S'_2 and S', respectively. The successor state axiom for $\textsc{InR}_1$ causes $\textsc{InR}_1$ to hold in these situations. Therefore, the agent believes $\textsc{InR}_1$ in S'. By the successor state axiom for $\textsc{Light}_1$, which we state below, the truth value of $\textsc{Light}_1$ would not change as the result of $\textsc{leave}$. This is an example of *belief update*: the agent's beliefs are modified as a result of reasoning about actions performed in the environment.

Now the agent performs the sensing action $\textsc{senseLight}$. According to the sensed fluent axioms for $\textsc{senseLight}$, $SF(\textsc{senseLight}, S^*)$ holds for situation S^* if and only if the light is on in the room in which the agent is located in S^*. In the figure, the light in R_1 is on in S' and S'_2, but not in S'_1. So, SF holds for $\textsc{senseLight}$ in the former two situations but not in the latter. The successor state axiom for B ensures that after doing a sensing action A, any situation that disagrees

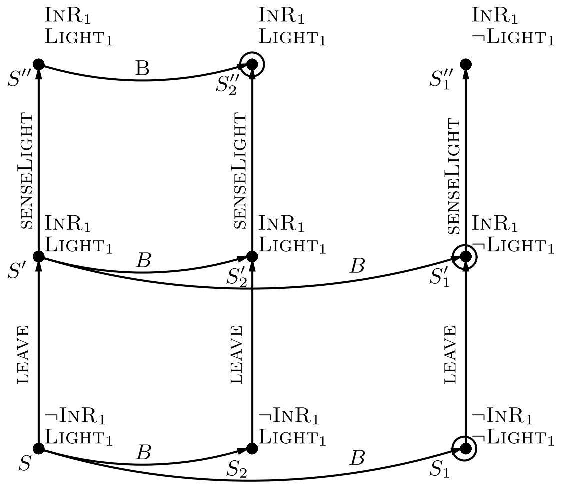

Fig. 2. An example of belief update and revision.

with the actual situation on the value of *SF* for *A* is dropped from the *B* relation in the successor state. In the figure, S' is the actual situation. Since S'_1 disagrees with S' on the value of *SF* for senseLight, $do(\text{senseLight}, S'_1)$ (labelled S''_1 in the figure) *is not* B-related to $do(\text{senseLight}, S')$ (labelled S''). On the other hand, S'_2 and S' agree on the value of *SF* for senseLight, so $do(\text{senseLight}, S'_2)$ (labelled S''_2 in the figure) *is* B-related to S''. The result is that the agent believes the light is on in S''. This is an example of belief expansion because the belief that the light is on was simply added to the belief state of the agent. *Belief revision* works using the same principles.

Our definition of *Bel* is similar to the one in Scherl and Levesque, but we are going to generalize their account in order to be able to talk about how *plausible* the agent considers a situation to be. Plausibility is assigned to situations using a function $pl(s)$, whose range is the natural numbers, where lower values indicate higher plausibility. The *pl* function only has to be specified over initial situations, using an initial state axiom. Successor situations have the same plausibility as their predecessors, as stipulated by the following successor state axiom:

Axiom 9 *(Successor state axiom for pl).*

$$pl\big(do(a, s)\big) = pl(s).$$

We say that the agent believes a proposition ϕ in situation s, if ϕ holds in the *most plausible* B-related situations. A situation is *most plausible* in situation s, if it is at least as plausible as all alternate situations to s:

Definition 10.

$$MP\big(s', s\big) \stackrel{\text{def}}{=} \forall s''.B\big(s'', s\big) \supset pl\big(s'\big) \leqslant pl\big(s''\big).$$

We use $MPB(s', s)$ to denote the situations s' that are most plausible and B-related to s:

Definition 11.

$$MPB\big(s', s\big) \stackrel{\text{def}}{=} B\big(s', s\big) \wedge MP\big(s', s\big).$$

Finally, we define the belief operator as follows:

Definition 12.

$$Bel(\phi, s) \stackrel{\text{def}}{=} \forall s'.MPB\big(s', s\big) \supset \phi\big[s'\big].$$

That is, ϕ is believed at s when it holds at all the most plausible situations B-related to s. Note that unlike Spohn [20] and Darwiche and Pearl [18], we do not require some situations to have plausibility 0.

We now return to the initial situations in Fig. 2, and add a plausibility structure to the belief state of the agent by supposing that S_1 is more plausible than S_2 (indicated by the circle surrounding S_1). For example, suppose that $pl(S_1) = 0$ and $pl(S_2) = 1$. Now, the beliefs of the agent are initially determined only by S_1. Therefore, the agent now believes that the light R_1 is off in S. After leaving R_2, the agent continues to believe that the light is off. After doing senseLight, S''_1 is dropped from B as before, so now S_2'' is the most plausible accessible situation, which means that it determines the beliefs of the agent. Since the light is on in S_2'', the agent believes it is on in S''. Since the agent goes from believing the light is off to believing it is on, this is a case of belief revision.

In order to ensure positive and negative introspection of beliefs, we assert the (TE) constraint over initial situations using an initial state axiom:

Axiom 13.

$$Init(s) \wedge B(s', s) \supset (\forall s''.B(s'', s') \equiv B(s'', s)).$$

The successor state axiom for B preserves this constraint over all situations.

Theorem 14.

$$\{\text{Axioms 1, 6, and 13}\} \models \forall s, s'.B(s', s) \supset (\forall s''.B(s'', s') \equiv B(s'', s)).$$

In order to clarify how this constraint ensures that introspection is handled properly, we will show that in the example illustrated in Fig. 2, the agent introspects its past beliefs. First, we need some notation that allows us to talk about the past. We use $\mathbf{Prev}(\phi, s)$ to denote that ϕ held in the situation immediately before s:

Definition 15.

$$\mathbf{Prev}(\phi, s) \stackrel{\text{def}}{=} \exists a, s'.s = do(a, s') \wedge \phi[s'].$$

Recall that in the example, the agent believed that the light in R_1 was off in S', i.e., $Bel(\neg\textsc{Light}_1, S')$. We want to show that $Bel(\mathbf{Prev}(Bel(\neg\textsc{Light}_1)), S'')$ holds, i.e., in S'', the agent believes that in the previous situation it believed that the light in R_1 was off. Consider a situation S^* that is among the most plausible B-related situations to S''. In this example, there is only one such situation, namely, S_2''. We need to show that $\mathbf{Prev}(Bel(\neg\textsc{Light}_1), S_2'')$ holds, i.e., that $Bel(\neg\textsc{Light}_1, S_2')$ holds. By Theorem 14, S_2' is B-related to the same situations as S', i.e., S_1' and S_2'. Since S_1' is more plausible than S_2', we only require that $\neg\textsc{Light}_1(S_1')$ holds. Since this is true, it follows that $Bel(\mathbf{Prev}(Bel(\neg\textsc{Light}_1)), S'')$ is also true.

The specification of *pl* and B over the initial situations is the responsibility of the axiomatizer of the domain in question. This specification need not be complete. Of course, a more complete specification will yield more interesting properties about the agent's current and future belief states.

We have another constraint on the specification of B over the initial situations: the situations B-related to an initial situation are themselves initial, i.e., the agent believes that initially nothing has happened. We assert this constraint as an initial state axiom:

Axiom 16.

$$Init(s) \wedge B(s', s) \supset Init(s').$$

3. Properties

In this section, we highlight some of the more interesting properties of our framework. In order to clarify our explanations and facilitate a comparison with previous approaches to belief change, it will be important for us to attach a specific meaning to the use of the terms *revision* and *update*, which we will do here. Let Σ denote the set of axioms of the previous sections (i.e., Axioms 1–16).

3.1. Belief revision

Recall from Section 2 that belief revision is suited to the acquisition of information about static environments for which the agent may have mistaken or partial information. In our framework, this can only be achieved through the use of sensing actions since they do not act to modify the environment but rather to tell us something about it. We suppose that for each formula ϕ by which we might want to revise, there is a corresponding sensing action capable of determining the truth value of ϕ. Moreover, we assume that this sensing action has no effect on the environment; the only fluent it changes is B.[11]

Definition 17 *(Uniform formula).* We call a formula *uniform* if the only situation term it contains is the situation constant *now* and it contains no unbound variables.

[11] This is not an overly strict imposition for we can capture sensing actions that modify the domain by "decomposing" the action into a sequence of non-sensing actions and sensing actions.

We now define a revision action as follows:

Definition 18 *(Revision action for ϕ).* A revision action A for a uniform formula ϕ with respect to action theory Σ is a sensing action that satisfies the following condition for every domain-dependent fluent F:

$$\Sigma \models \left[\forall s. SF(A,s) \equiv \phi[s]\right] \wedge \left[\forall s \forall \vec{x}. F(\vec{x},s) \equiv F\left(\vec{x}, do(A,s)\right)\right].$$

In other words, A is a sensing action for the formula ϕ, and it does not change any physical fluents. Since we assume there is a revision action A for each formula ϕ that we might want to revise by, we assume that Σ also contains the appropriate sensed fluent axioms and successor state axioms to satisfy this definition.

Definition 19 *(Domain-dependent formula).* We refer to a formula as *domain-dependent* if all the fluents mentioned in it are domain-dependent.

It is easy to see from Definition 18 that if A is a revision action and ϕ^* is domain-dependent, then A does not affect the value of ϕ^*.

Lemma 20. *Let ϕ^* be a domain-dependent formula, and A be a revision action for some formula ϕ. Then*:

$$\Sigma \models \forall s. \phi^*[s] \equiv \phi^*\left[do(A,s)\right].$$

We now show that belief revision is handled appropriately in our system in the sense that if the sensor indicates that ϕ holds, then the agent will indeed believe ϕ after performing A. Similarly, if the sensor indicates that ϕ is false, then the agent will believe $\neg\phi$ after doing A.

Theorem 21. *Let ϕ be a domain-dependent, uniform formula, and A be a revision action for ϕ with respect to Σ. It follows that*:

$$\Sigma \models \left[\forall s. \phi[s] \supset Bel\left(\phi, do(A,s)\right)\right] \wedge \left[\forall s. \neg\phi[s] \supset Bel\left(\neg\phi, do(A,s)\right)\right].$$

If the agent is indifferent towards ϕ before doing the action, i.e., does not believe ϕ or $\neg\phi$, this is a case of belief expansion. If, before sensing, the agent believes the opposite of what the sensor indicates, then we have belief revision.

Note that this theorem also follows from Scherl and Levesque's theory. However, for Scherl and Levesque, if the agent believes $\neg\phi$ in S and the sensor indicates that ϕ is true, then in $do(A,S)$, the agent's belief state will be inconsistent. The agent will then believe all propositions, including ϕ. In our theory, the agent's belief state will be consistent in this case, as long as there is some situation S' accessible from S that agrees with S on the value of the sensor associated with A.

Theorem 22. *Let A be a revision action for a domain-dependent, uniform formula ϕ with respect to Σ. The following set of sentences (which we denote by Γ) is satisfiable*:

$$\Sigma \cup \left\{Bel(\neg\phi, S_0), Bel\left(\phi, do(A, S_0)\right), \neg Bel\left(FALSE, do(A, S_0)\right)\right\}.$$

3.2. Belief update

Belief update refers to the belief change that takes place due to a change in the environment. In analogy to revision, we introduce the notion of an update action.

Definition 23 *(Update action for ϕ).* An update action A for a uniform formula ϕ with respect to action theory Σ is a non-sensing action that always makes ϕ true in the environment. That is, $\Sigma \models \forall s. \phi[do(A,s)] \wedge SF(A,s)$.

As with Scherl and Levesque's theory, the agent's beliefs are updated appropriately when an update action A for ϕ occurs, i.e., the agent will believe ϕ after A is performed.

Theorem 24. *Let A be an update action for ϕ. Then*:

$$\Sigma \models \forall s. Bel\left(\phi, do(A,\ s)\right).$$

In our framework, we can represent actions that do not fall under the category of update actions. Of particular interest are ones whose effects depend on what is true in the current situation, i.e., conditional effects. We can prove an analogous theorem for such actions. Suppose that A is a non-sensing action, i.e., $\Sigma \models \forall s. SF(A,s)$, and that A is an action that causes ϕ' to hold, whenever ϕ holds beforehand. Further suppose that the agent believes ϕ in S. Then, after performing A in S, the agent ought to believe that ϕ' holds.

Theorem 25. *Let A be a ground action term, and ϕ, ϕ' be uniform formulae. Then*:

$$\Sigma \models \forall s.Bel(\phi, s) \wedge \forall s' SF(A, s') \wedge (\forall s'.\phi[s'] \supset \phi'[do(A, s')]) \supset Bel(\phi', do(A, s)).$$

It is very important to note that in our framework, there are no actions that correspond directly to the actions "revise by ϕ" or "update by ϕ". We only have physical actions and sensing actions. It is, therefore, the properties associated with these actions by the successor state and sensed fluent axioms that determine how (and whether) the agent's beliefs get revised or updated.

3.3. Introspection

Since we constrained the accessibility relation to be transitive and Euclidean, our agents are guaranteed to be introspective.

Theorem 26.

$$\Sigma \models [Bel(\phi, s) \supset Bel(Bel(\phi), s)] \wedge [\neg Bel(\phi, s) \supset Bel(\neg Bel(\phi), s)].$$

3.4. Awareness of mistakes

In Section 2.2, we claimed that the agent can also introspect its past beliefs. Suppose that the agent believes $\neg\phi$ in S, and after performing a revision action A for ϕ in S, the agent believes ϕ. In $do(A, S)$, the agent should also believe that in the previous situation ϕ was true, but it believed ϕ was false. In other words, the agent should believe that it was mistaken about ϕ. We now prove a theorem that states that the agent will indeed believe that it was mistaken about ϕ.

Theorem 27. *Let A be a revision action for a domain-dependent, uniform formula ϕ with respect to Σ. Then*:

$$\Sigma \models \forall s.Bel(\neg\phi, s) \wedge Bel(\phi, do(A, s)) \supset Bel(\mathbf{Prev}(\phi \wedge Bel(\neg\phi)), do(A, s)).$$

The properties presented in this section demonstrate the elegance and the power of our framework. While the framework itself is not overly complex, it provides a powerful system in which to reason about the beliefs of an agent in a dynamic environment with the capability to perform actions and to sense the environment.

4. Example

We now present an example to illustrate how this theory of belief change can be applied. We model a world in which there are two rooms, R_1 and R_2. The agent can move between the rooms. Each room contains a light that can be on or off. The agent has two binary sensors. One sensor detects whether or not the light is on in the room in which the agent is currently located. The other sensor detects whether or not the agent is in R_1.

We have three fluents: $\textsc{Light}_1(s)$ ($\textsc{Light}_2(s)$, respectively), which holds if and only if there is light in R_1 (R_2, respectively) in situation s, and $\textsc{InR}_1(s)$, which holds if the agent is in R_1 in s. If the agent is not in R_1, then it is assumed to be in R_2. There are three actions: the agent leaves the room it is in and enters the other room ($\textsc{leave}$), the agent senses whether it is in R_1 ($\textsc{senseInR}_1$), and the agent senses whether the light is on in the room in which it is currently located ($\textsc{senseLight}$).

The successor state axioms and guarded sensed fluent axioms for our example, which we will call E, are as follows:

$$\textsc{Light}_1(do(a, s)) \equiv \textsc{Light}_1(s)$$
$$\textsc{Light}_2(do(a, s)) \equiv \textsc{Light}_2(s)$$
$$\textsc{InR}_1(do(a, s)) \equiv ((\neg\textsc{InR}_1(s) \wedge a = \textsc{leave}) \vee (\textsc{InR}_1(s) \wedge a \neq \textsc{leave}))$$
$$TRUE \supset (SF(\textsc{leave}, s) \equiv TRUE)$$
$$\textsc{InR}_1(s) \supset (SF(\textsc{senseLight}, s) \equiv \textsc{Light}_1(s))$$
$$\neg\textsc{InR}_1(s) \supset (SF(\textsc{senseLight}, s) \equiv \textsc{Light}_2(s))$$
$$TRUE \supset (SF(\textsc{senseInR}_1, s) \equiv \textsc{InR}_1(s))$$

Next we must specify the initial state. This includes both the physical state of the domain and the belief state of the agent. First we describe the initial physical state of the domain, by saying which domain-dependent fluents hold in the actual initial situation, S_0. Initially, the lights in both rooms are on and the agent is in R_2 (this is illustrated on the left-hand side of Fig. 3):

$$\textsc{Light}_1(S_0) \wedge \neg\textsc{InR}_1(S_0) \wedge \textsc{Light}_2(S_0).$$

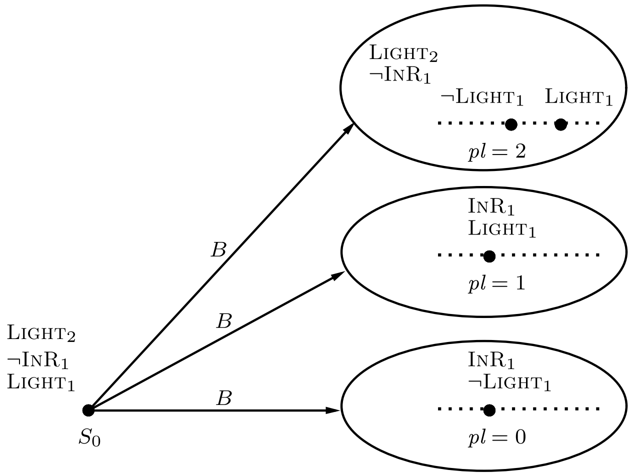

Fig. 3. The initial state of the example domain.

The initial belief state of the agent is illustrated in Fig. 3. It shows that in the most plausible situations B-related to S_0 (the ones with plausibility 0 in the figure), $\neg\mathrm{Light}_1$ and InR_1 hold. In the next most plausible situations B-related to S_0 (the ones with plausibility 1), Light_1 and InR_1 hold. In the third most plausible (the ones with plausibility 2) B-related situations to S_0, Light_2 and $\neg\mathrm{InR}_1$ hold. There is also at least one situation in the latter group in which Light_1 holds and one in which $\neg\mathrm{Light}_1$ holds. Specifying this belief state directly can be cumbersome. For example, the axiom for the situations with plausibility 1 is:

$$\big(\exists s.\mathit{Init}(s) \wedge B(s, S_0) \wedge pl(s) = 1\big) \wedge$$
$$\big(\forall s.\mathit{Init}(s) \wedge pl(s) = 1 \supset \mathrm{Light}_1(s) \wedge \mathrm{InR}_1(s)\big).$$

For now, we will not enumerate the set of axioms that specify the belief state shown in Fig. 3. But we assume that we have such a set which, together with the axioms for the initial physical state, we refer to as I. After we have discussed the example, we will show that there is a more elegant way to specify the initial belief state of the agent. So for this example, we add E, and I to Σ, and obtain the following:

Proposition 28. *The following formulae are entailed by* $\Sigma \cup E \cup I$:

1. $\mathit{Bel}(\neg\mathrm{Light}_1 \wedge \mathrm{InR}_1, S_0)$
2. $\mathit{Bel}(\mathrm{Light}_1 \wedge \mathrm{InR}_1, \mathit{do}(\mathrm{senseLight}, S_0))$
3. $\mathit{Bel}(\neg\mathrm{InR}_1, \mathit{do}(\mathrm{senseInR}_1, \mathit{do}(\mathrm{senseLight}, S_0)))$
4. $\mathit{Bel}(\mathbf{Prev}(\neg\mathrm{InR}_1 \wedge \mathit{Bel}(\mathrm{InR}_1)), \mathit{do}(\mathrm{senseInR}_1, \mathit{do}(\mathrm{senseLight}, S_0)))$
5. $\neg\mathit{Bel}(\mathrm{Light}_1, \mathit{do}(\mathrm{senseInR}_1, \mathit{do}(\mathrm{senseLight}, S_0))) \wedge$
   $\neg\mathit{Bel}(\neg\mathrm{Light}_1, \mathit{do}(\mathrm{senseInR}_1, \mathit{do}(\mathrm{senseLight}, S_0)))$
6. $\mathit{Bel}(\mathrm{InR}_1, \mathit{do}(\mathrm{leave}, \mathit{do}(\mathrm{senseInR}_1, \mathit{do}(\mathrm{senseLight}, S_0))))$
7. $\mathit{Bel}(\mathrm{Light}_1,$
   $\mathit{do}(\mathrm{senseLight},$
   $\mathit{do}(\mathrm{leave}, \mathit{do}(\mathrm{senseInR}_1,$
   $\mathit{do}(\mathrm{senseLight}, S_0))))).$

We shall now give a short, informal explanation of why each part of the previous theorem holds.

1. In the most plausible situations B-related to S_0, $\neg\mathrm{Light}_1 \wedge \mathrm{InR}_1$ holds.
2. Even though the agent believes that it is in R_1 initially, it is actually in R_2. Therefore, its light sensor is measuring whether there is light in R_2, even though the agent thinks that it is measuring whether there is light in R_1. It turns out that there is light in R_2 in S_0, so the sensor returns 1. Since the agent believes that the light sensor is measuring whether there is light in R_1 and in all the situations with plausibility 0, there is no light in R_1, those situations are dropped from the B relation. In the situations with plausibility 1, the light is on in R_1, so those situations are retained. In those situations $\mathrm{Light}_1 \wedge \mathrm{InR}_1$ holds and those fluents are not affected by the $\mathrm{senseLight}$ action, so the agent believes $\mathrm{Light}_1 \wedge \mathrm{InR}_1$ after doing $\mathrm{senseLight}$.
3. Now the agent senses whether it is in R_1. Again the agent's most plausible situations conflict with what is actually the case, so they are dropped from the B relation. The situations with plausibility 2 become the most plausible situations,

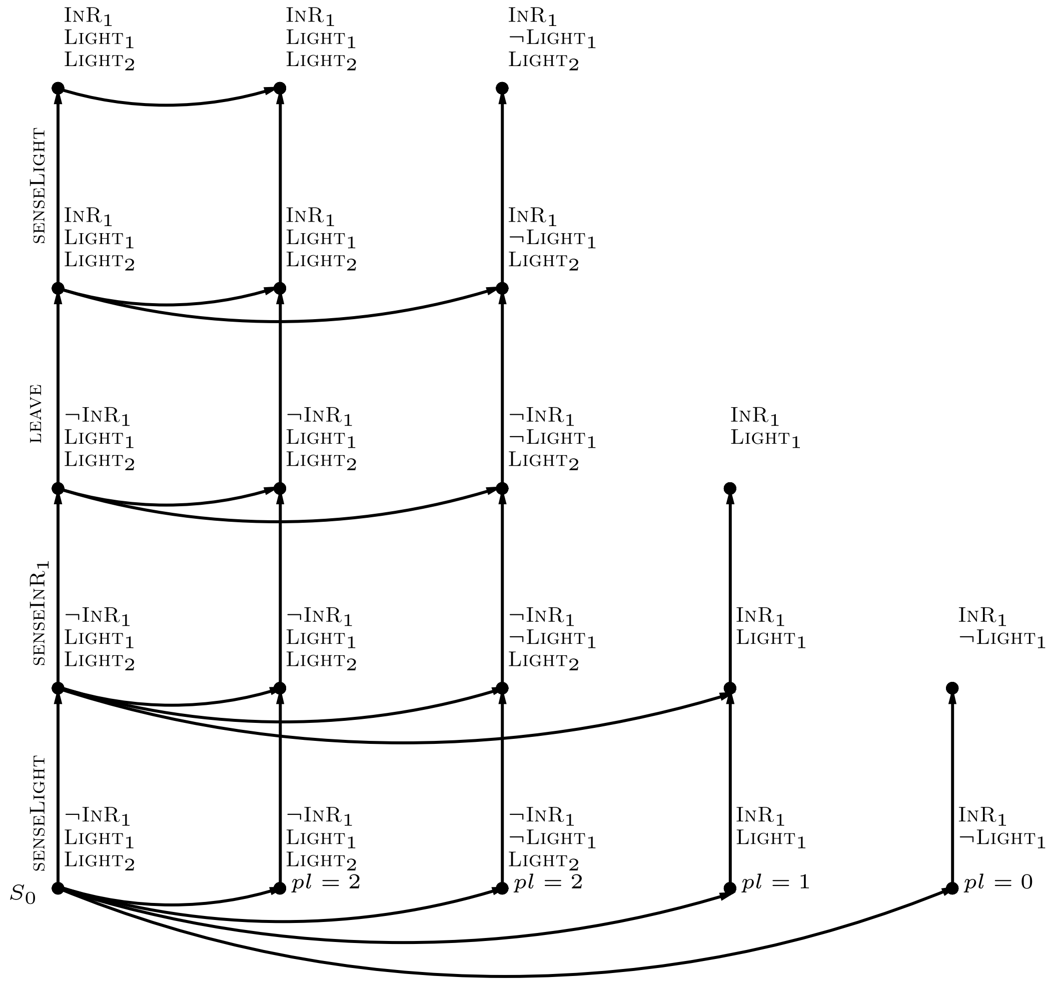

Fig. 4. Example domain. Only the important details are displayed.

so the agent believes it is not in R_1. By using plausibilities we allow for "fallback" situations and thus deal with settings that Scherl and Levesque [5,6] cannot handle. Their formalism would have descended into inconsistency in this case as there would be no situations in which $\neg$InR$_1$ is possible. In fact, this would already have happened after the first sensing action senseLight. Scherl and Levesque can only reason about situations that are consistent with the agent's current beliefs and therefore cannot deal with sensing results that conflict with these beliefs where an account of belief revision, such as the one offered here, is required.

4. By Theorem 27, the agent realizes that it was mistaken about being in R_1.
5. Among the situations with plausibility 2, there is one in which the light is on in R_1 and one in which it is not on. Therefore, the agent is unsure as to whether the light is on.
6. Now the agent leaves R_2 and enters R_1. This happens in all the B-related situations as well. Therefore, the agent believes that it is in R_1. This is an example of an update.
7. The light in R_1 was on initially, and since no action was performed that changed the state of the light, the light remains on. After checking its light sensor, the agent believes that the light is on in R_1.

A more complete illustration of this example is given in Fig. 4.

This example shows that the agent's beliefs change appropriately after both revision actions and update actions. The example also demonstrates that our formalism can accommodate iterated belief change. The agent goes from believing that the light is not on, to believing that it is on, to not believing one way or the other, and then back to believing that it is on.

To facilitate the specification of the initial belief state of the agent, we find it convenient to define another belief operator $\Rightarrow$, in the spirit of the conditional logic connective [21]:

Definition 29.

$$\phi \Rightarrow_s \psi \stackrel{\text{def}}{=}$$
$$\forall s'.B(s', s) \wedge \phi[s'] \wedge (\forall s''.B(s'', s) \wedge \phi[s''] \supset pl(s') \leqslant pl(s'')) \supset \psi[s'].$$

$\phi \Rightarrow_s \psi$ holds if in the most plausible situations B-related to s where ϕ holds, ψ also holds. Note that for any situation S, $Bel(\phi, S)$ is equivalent to $(TRUE \Rightarrow_S \phi)$.

We can use this operator to specify the initial belief state of the agent without having to explicitly mention the plausibility of situations. To obtain the results of Proposition 28, it suffices to let I be the following set of axioms:

$$\textsc{Light}_1(S_0) \wedge \neg \textsc{InR}_1(S_0) \wedge \textsc{Light}_2(S_0)$$

$$TRUE \Rightarrow_{S_0} \neg \textsc{Light}_1 \wedge \textsc{InR}_1$$

$$\textsc{Light}_1 \Rightarrow_{S_0} \textsc{InR}_1$$

$$\neg(\textsc{Light}_2 \wedge \neg \textsc{InR}_1 \Rightarrow_{S_0} \textsc{Light}_1)$$

$$\neg(\textsc{Light}_2 \wedge \neg \textsc{InR}_1 \Rightarrow_{S_0} \neg \textsc{Light}_1)$$

It is easy to see that the belief state depicted in Fig. 3 satisfies these axioms. In the most plausible worlds, $(\neg \textsc{Light}_1 \wedge \textsc{InR}_1)$ holds. In the most plausible worlds where the light in R_1 is on, the agent is in R_1. Finally, the last two axioms state that among the most plausible worlds where the light is on in R_2 and the agent is in R_2, there is one where the light is off in R_1 and one in which the light is on (respectively).

5. Postulate soundness

In this section, we consider the extent to which our framework satisfies the AGM postulates for belief revision, the KM postulates for belief update, and the DP postulates for iterated belief revision. In order to do so we first need to establish a common footing. The first notion to establish is what is meant by the belief state of the agent. We define a belief state (relative to a given situation) to consist of those formulae believed true at a particular situation. We limit our attention to uniform, domain-dependent formulae, since these frameworks only consider objective formulae (i.e., formulae without belief operators), and they do not have an explicit representation of state, but rather they implicitly refer to the current state, and so there is no need to consider beliefs regarding more than one situation. Therefore, the language we use here, $\mathcal{L}_{now}$, is a set of domain-dependent uniform formulae. We assume that $\mathcal{L}_{now}$ is propositional and finite.[12] ϕ, ψ, and γ will be used to denote domain-dependent uniform formulae. Also, t and u, possibly decorated, will be used to denote ground situation terms.

In the previous section, we used a theory to specify the beliefs of the agent. We said that the agent believed (did not believe, respectively) a formula, if the theory entailed that the formula was believed (not believed, respectively). Unless the theory is complete with respect to the beliefs of the agent, we will not be able to determine whether some formulae are believed or not. In other words, there could be more than one model of the theory. This is not the case for the semantic frameworks for belief change, e.g., the AGM framework. There the belief state of the agent is also determined by a set of sentences, however, there is also an implicit closed world assumption. The sentences in the set are believed by the agent, and if a sentence is not in the set, then it is not believed by the agent. There is no uncertainty about what is believed by the agent, i.e., the belief state of the agent can be represented by a single model. To bring our framework in line with the AGM framework, we assume that we have a model M of Σ, which will be used to fix the belief state of the agent. We can encode an AGM belief state K in our setting using M and a ground situation term t such that $M \models Bel(\phi, t)$ if and only if $\phi \in K$.

We need to define the three operators used in the postulates: belief expansion, revision and update. The operators in the postulates map belief states to new belief states. Our framework is based on situations and actions rather than belief sets and their operators, so we will use a different representation and then translate the postulates appropriately. Our revision and update operators are actions, i.e., they map situations into situations. Sensing actions lead to belief revision, and physical actions yield belief update. Iterated revisions and updates are handled using sequences of actions. However, there is no action type that corresponds to belief expansion. Therefore, we define the expansion operator directly as a belief state as discussed below. As a consequence, expansions cannot be iterated, but the postulates do not require that we iterate them.

We first define a function that maps a situation t into the belief state of the agent at t.

12 We make this assumption to accord with the AGM and KM frameworks, however this assumption is only used in the proof for the soundness of $(\mathrm{K} \diamond 8)$ (Lemma 57).

Definition 30 *($K(t)$).* We denote the belief state at t (in M) by $K(t)$ and define it as follows:

$$K(t) = \{\psi\colon\ M \models Bel(\psi, t)\}.$$

It is easily verified that $K(t)$ is closed under deduction.

Recall that the expansion of a belief state K by a formula ϕ is defined to be the belief state that results from simply adding ϕ to K. If ϕ is inconsistent with K, then the belief state that results from adding ϕ to K will be inconsistent. We can encode this in our framework as an operator that maps a situation t and a formula ϕ into a belief state, but this time it will be the belief set that results from adding ϕ to $K(t)$.

Definition 31 *($t + \phi$).* We denote the expansion of t by ϕ (in M) by $t + \phi$ and define it as follows:

$$t + \phi = \{\psi\colon\ M \models Bel(\phi \supset \psi, t)\}.$$

So, the belief state that results from expanding t with ϕ is the set of formulae that are believed to be implied by ϕ in t. Since our expansion operator does not return a situation, it can only be applied last in a sequence of operations. This is the case for all of the postulates, so this definition suffices for our purposes.

Next, we define the revision of t by ϕ, $t * \phi$, as the situation that results from performing a revision action for ϕ in t. As we said earlier, we assume that for any ϕ under consideration, there is a revision action A_ϕ for ϕ with respect to Σ. When ϕ is clear from the context, we will drop the subscript. In the AGM setting, a revision $K * \phi$ is interpreted as the revision of beliefs K *after learning* ϕ. In our case, we do not know whether ϕ will be true until after performing the revision action. Accordingly, we define a revision of t by ϕ only in the case that ϕ happens to be true in situation t (i.e., $M \models \phi[t]$).

Definition 32 *($t * \phi$).* We denote the revision of t by ϕ (in M) as $t * \phi$ and define it as follows:

$$t * \phi = do(A_\phi, t),$$

whenever $M \models \phi[t]$. If $M \models \neg\phi[t]$, then $t * \phi$ is undefined.

5.1. AGM postulates

Here are the AGM postulates translated into our notation.

(K*1) $K(t * \phi)$ is deductively closed
(K*2) $\phi \in K(t * \phi)$
(K*3) $K(t * \phi) \subseteq t + \phi$
(K*4) If $\neg\phi \notin K(t)$, then $t + \phi \subseteq K(t * \phi)$
(K*5) $K(t * \phi) \neq \mathcal{L}_{now}$
(K*6) If $\models \phi \equiv \psi$, then $K(t * \phi) = K(t * \psi)$
(K*7) $K(t * \phi \wedge \psi) \subseteq (t * \phi) + \psi$
(K*8) If $\neg\psi \notin K(t * \phi)$, then $(t * \phi) + \psi \subseteq K(t * \phi \wedge \psi)$

Note that (K*5) is somewhat different from the corresponding standard AGM postulate. This is due to the fact that it is not possible in our framework to revise with an identically false formula, since $t * \phi$ is only defined if $M \models \phi[t]$. To obtain (K*5) in our framework, we require a further assumption, i.e., that initially the agent does not think that the actual situation is completely implausible, i.e., B is reflexive.[13] This does not mean that we get knowledge instead of belief, since the actual situation does not have to be most plausible, but it must not be completely implausible. We use an initial state axiom to state this assumption:

Axiom 33.

$$Init(s) \supset B(s, s).$$

We add this axiom to Σ, and it follows that B is everywhere reflexive:

Theorem 34.

$$\Sigma \models \forall s B(s, s).$$

[13] Note that strictly speaking, we only need the initial situation that precedes t to be self-accessible, however the theorem is easier to state if we assume that all initial situations are self-accessible.

As an immediate consequence, we have that the agent's beliefs never become inconsistent.

Corollary 35.

$$\Sigma \models \forall s \neg Bel(FALSE, s).$$

Finally, we are now able to show that our translations of the AGM postulates are satisfied.

Theorem 36. *For any model M of Σ, the AGM postulates* (K*1)–(K*8) *are satisfied, when $*$ is defined as in Definition* 32 *for any situation t and domain-dependent uniform formulae ϕ and ψ.*

5.2. KM postulates

We now turn to defining belief update in our framework, which is defined analogously to revision but using an update action instead of a revision action. The update of a situation t with a *consistent* formula ϕ is the situation that results from performing an update action for ϕ in t. We can only handle consistent formulae because according to the definition of an update action A, $\Sigma \models \forall s.\phi[do(A, s)]$, which must be false if ϕ is inconsistent. We will assume that for any consistent ϕ under consideration, there is at least one update action for ϕ with respect to Σ, and that we have a function *ua* which maps a consistent formula ϕ into an update action for ϕ such that for any ψ, if $\models \phi \equiv \psi$ then $ua(\phi) = ua(\psi)$ (we only need this condition for postulate $(\mathrm{K} \diamond 5)$).

Definition 37. We define the update of t by ϕ to be:

$$t \diamond \phi = do\big(ua(\phi), t\big).$$

In order to translate postulate $(\mathrm{K} \diamond 8)$ into our framework, we need to define $\bigcap_{w \in [K]} w \diamond \phi$. As we saw above, $[K]$ is the set of complete, consistent (cc) theories that contain K, and a cc theory can be thought of as a possible world. We use situations as possible worlds, so given a belief set at a situation t, we need a set of situations that corresponds to the set of all cc extensions of $K(t)$. We can easily map situations into cc theories.

Definition 38 *($Tr(t)$).* We define the truths at t (in M) to be:

$$Tr(t) = \big\{\psi\colon\ M \models \psi[t]\big\}.$$

Since $\mathcal{L}_{now}$ is propositional, the minimal accessible situations from t coincide with the cc extensions of $K(t)$. Let $MPB(t)$ denote $\{t'\colon\ MPB(t', t)\}$.

Lemma 39. *$w \in [K(t)]$ if and only if there exists $t' \in MPB(t)$ such that $w = Tr(t')$.*

Now, we can translate the equation in the postulate as follows:

$$K(t \diamond \phi) = \bigcap_{t' \in MPB(t)} Tr\big(t' \diamond \phi\big).$$

In other words, the belief set that results from updating t by ϕ is the same as the intersection of the truths that result from updating each member of $MPB(t)$ by ϕ.

Here is a translation of the KM postulates into our notation[14]:

$(\mathrm{K} \diamond 1)$ $K(t \diamond \phi)$ is deductively closed
$(\mathrm{K} \diamond 2)$ $\phi \in K(t \diamond \phi)$
$(\mathrm{K} \diamond 3)$ If $\phi \in K(t)$, then $K(t \diamond \phi) = K(t)$
$(\mathrm{K} \diamond 4)$ $K(t \diamond \phi) = \mathcal{L}_{now}$ iff $K(t) \models FALSE$
$(\mathrm{K} \diamond 5)$ If $\models \phi \equiv \psi$, then $K(t \diamond \phi) = K(t \diamond \psi)$
$(\mathrm{K} \diamond 6)$ $K(t \diamond (\phi \wedge \psi)) \subseteq (t \diamond \phi) + \psi$
$(\mathrm{K} \diamond 7)$ If $K(t)$ is complete and $\neg\psi \notin K(t \diamond \phi)$, then $(t \diamond \phi) + \psi \subseteq K(t \diamond (\phi \wedge \psi))$
$(\mathrm{K} \diamond 8)$ If $MPB(t) \neq \emptyset$ then $K(t \diamond \phi) = \bigcap_{t' \in MPB(t)} Tr(t' \diamond \phi)$

Theorem 40. *$K(t \diamond \phi)$ satisfies KM postulates* $(\mathrm{K} \diamond 1)$, $(\mathrm{K} \diamond 2)$, $(\mathrm{K} \diamond 4)$, $(\mathrm{K} \diamond 5)$, $(\mathrm{K} \diamond 8)$ *when $\diamond$ is defined as in Definition* 37 *for any situation t and consistent domain-dependent uniform formulae ϕ and ψ.*

[14] For $(\mathrm{K} \diamond 4)$, note that if $\phi \models FALSE$, then $t \diamond \phi$ is not defined, since ϕ has to be consistent.

Notice that postulates (K $\diamond$ 3), (K $\diamond$ 6), and (K $\diamond$ 7) are not satisfied because while an update action for ϕ is guaranteed to make ϕ true, it may have other effects. In fact, it is not possible in our framework to define an update action that makes an arbitrary ϕ true and no other changes because our action theories do not handle disjunctive effects of actions. If ϕ is of the form $\psi_1 \vee \psi_2$, then we could have an action that only makes ψ_1 true and one that only makes ψ_2 true, but not one that only makes the disjunction true.

Boutilier [19] has a problem with (K $\diamond$ 3) ((U2) in the KM rendering) for similar reasons. In his framework, (update) actions have plausibilities, and the most plausible action explaining the new information is assumed to have taken place. It could be that this action has other effects. To satisfy this postulate, he introduces a *null event* and considers a model in which this is the most plausible event at any world.

5.3. DP postulates

In our framework, iterated revision corresponds to the performing of at least two consecutive revision actions. We now show that there is some correspondence with the Darwiche and Pearl [18] account of iterated belief revision. Here are the DP postulates translated into our notation.

(DP1) If $\psi \models \phi$, then $K((t*\phi)*\psi) = K(t*\psi)$
(DP2) If $\psi \models \neg\phi$, then $K((t*\phi)*\psi) = K(t*\psi)$
(DP3) If $\phi \in K(t*\psi)$, then $\phi \in K((t*\phi)*\psi)$
(DP4) If $\neg\phi \notin K(t*\psi)$, then $\neg\phi \notin K((t*\phi)*\psi)$

Theorem 41. *Postulates* (DP1), (DP3) *and* (DP4) *are satisfied when $*$ is defined as in Definition* 32 *for any situation t and domain-dependent uniform formulae ϕ and ψ.*

Interestingly, changes of the type described by (DP2) are not defined according to our view of belief revision. In the case where $\psi \models \neg\phi$, either $t*\phi$ or $t*\psi$ is undefined, therefore $K((t*\phi)*\psi)$ is undefined.

6. Discussion

There are various aspects of our framework that deserve further consideration. We address what we consider to be some of the more important issues here.

This work continues the tradition begun by John McCarthy [2,3,22] exploring the use of symbolic representations for representing and reasoning about dynamic systems. The legacy of the situation calculus has proved an influential and lasting contribution to research on reasoning about action and change. It provides the foundations upon which our framework is built. McCarthy's early work [2,3] supplied the basic theory underlying the situation calculus and it is from here that we take our departure point. His later work [22] identifies the *frame problem* and other issues that need to be addressed when using the situation calculus. While we have not utilized McCarthy's [23] circumscription to solve the frame problem but rather followed Reiter's [4] successor state axiom approach, Reiter's approach is certainly influenced by the development of circumscription and its use to solve the frame problem. Our approach is also built on the insights of Moore [11] who reifies the accessibility relation used in the semantics for modal logics. This approach to modality is supported by McCarthy [24].

Our plausibility function is based on ordinal conditional functions [18,20] (particularly Spohn's κ-rankings). However, our assignment of plausibilities to situations is fixed, whereas in most frameworks based on assigning plausibilities to worlds, the plausibility assigned to a world can change when revisions occur. The dynamics of belief in our framework derives from the dynamics of the B-relation, rather than that of the plausibility assignment. Note that Friedman and Halpern [25] make similar assumptions to ours. In Darwiche and Pearl's framework [18], the κ-ranking of a world that does not satisfy the formula in a revision increases by 1. However, if the world satisfies the revision formula in future revisions, the world's κ-ranking decreases, and if it decreases to 0, the world will take part in determining the beliefs of the agent. In our framework, when a sensing action occurs, any situation S' that disagrees with the actual value of the sensor is *removed* from the B relation (actually, its successor is removed). The successors of S' will never be readmitted to B, so they will never help determine the beliefs of the agent. This amounts to saying that the information that the agent learns from sensing is knowledge, i.e., the agent will never get new information that contradicts previous sensing information. However, this is quite reasonable since our framework assumes exact sensing and that there are no exogenous actions, so the agent should expect that its sensory information will not be contradicted. For a generalization of our framework to the noisy sensing case that also allows plausibility update, please see [26]. Our framework was generalized to handle exogenous actions in [27].

One may think that having a fixed plausibility assignment limits the applicability of our approach. Consider an example[15] where, most plausibly, a cat is asleep at home, but where after phoning home, most plausibly, the cat is awake. (Nothing is certain in either case.) This might seem to require adjustment of the plausibility assignment to situations.

[15] We are indebted to Jim Delgrande for this example.

To handle this example, we need first to observe that in the action theory we are using, actions are taken to be *deterministic*, with effects described by successor state axioms, quite apart from properties of belief and plausibility. If in some situations a phone action wakes the cat, and in others not, then there has to be some property M such that we can write a successor state axiom of the following form:

$$\begin{aligned}\textsc{Awake}\big(do(a,s)\big) \equiv {} & \big(a = \textsc{phone} \wedge M(s)\big) \vee \\ & [\ldots \text{other actions that can wake cats} \ldots] \vee \\ & \big(\textsc{Awake}(s) \wedge [a \text{ is not some put-to-sleep action}]\big).\end{aligned}$$

For example, M could represent that "the phone's ringer is loud enough to wake the cat". With this model, we can then arrange the B relation in the initial situation so that there are four groups of situations s' B-related to S_0, where the following hold (in order of decreasing plausibility): (1) $M(s') \wedge \neg\textsc{Awake}(s')$, (2) $M(s') \wedge \textsc{Awake}(s')$, (3) $\neg M(s') \wedge \neg\textsc{Awake}(s')$, and (4) $\neg M(s') \wedge \textsc{Awake}(s')$. Then we obtain that

$$Bel(\neg\textsc{Awake}, S_0),$$

holds since the most plausible situations s' that are B-related to S_0 satisfy $M(s') \wedge \neg\textsc{Awake}(s')$. However, the most plausible situations B-related to $do(\textsc{phone}, S_0)$ are those situations $do(\textsc{phone}, s')$ where s' is from group (1). Since $M(s')$ holds, so does $\textsc{Awake}(do(\textsc{phone}, s'))$ by the successor state axiom for Awake above. Therefore,

$$Bel\big(\textsc{Awake}, do(\textsc{phone}, S_0)\big),$$

holds, exactly as desired.[16] Of course, in this formalization, we also get that:

$$Bel(M, S_0),$$

but this is to be expected: why would we believe it most likely that the cat would be awake after the phone rings, if we did not also believe it most likely that the ringer was loud enough to waken it? In sum, we can account for changing our minds about the plausibility of the cat being awake without needing to change the plausibility ordering over situations.

We can also handle a variant of this example where we change our mind about whether phoning home wakes the cat. For example, imagine a sensing action $\textsc{examineRinger}$ that informs us that M is false initially (e.g., the ringer on the phone is set to low). Then, we get that

$$Bel\big(\neg\textsc{Awake}, do\big(\textsc{phone}, do(\textsc{examineRinger}, S_0)\big)\big)$$

holds, since the most plausible situations will now be descendants of the s' that are B-related to S_0 in group (3), where $\neg M(s') \wedge \neg\textsc{Awake}(s')$ holds. This is exactly as desired, and again without needing to change the plausibility assignment.

In the process of developing the approach described in this paper, we experimented with various schemes where the plausibility assigned to situations could be updated. But as discussed in Section 2.1 we found that this led to problems for introspection. Consider a scheme where we combine the plausibility assignment with the belief accessibility relation by adding an extra argument to the B relation, i.e., where $B(s', n, s)$ means that in situation s the agent thinks s' is plausible to degree n. In order to ensure that beliefs are properly introspected, the relation would have to satisfy the constraint (TE) discussed in Section 2.1, which is similar to the one given in Theorem 14, but taking plausibilities into account. That is to say, all the B-related situations to a situation s must have the same belief structure as s, i.e., they should be B-related to the same situations with the same plausibilities as s. Unfortunately, this conflicts with some of our intuitions about how to change plausibilities to accommodate new information.

Consider an example where we have two situations S_0 and S_1, and where initially the agent considers situation S_1 more plausible than S_0, i.e., $B(S_1, 0, S_0)$, $B(S_0, 1, S_0)$, $B(S_1, 0, S_1)$, $B(S_0, 1, S_1)$. Notice that S_0 and S_1 have the same belief structure. Suppose that $\textsc{Light}_1(S_0) \wedge SF(\textsc{senseLight}, S_0)$ holds as does $\neg\textsc{Light}_1(S_1) \wedge \neg SF(\textsc{senseLight}, S_1)$. The natural way to update the plausibilities after sensing would be to make the most plausible situations from a situation $do(\textsc{senseLight}, s)$ be the ones that agree with s on the value of $SF(\textsc{senseLight})$. So, if we let S'_0 denote $do(\textsc{senseLight}, S_0)$ and S'_1 denote $do(\textsc{senseLight}, S_1)$, then in S'_0, S'_0 should be more plausible than S'_1 and in S'_1, S'_1 should be more plausible than S'_0. But this would violate the constraint that B-related situations have the same belief structure, and cause introspection to fail.

One way to avoid this problem would be to update the plausibilities of all situations based on what holds in the 'actual' situations, i.e., S_0 and its successors (this focuses attention on beliefs that hold in actual situations, which is what we normally do anyway). For the example above, we would look at how the plausibilities should change in S'_0 and adjust the plausibilities in the situations B-related to S'_0 (in this case just S'_1) in the same way. We would then have that S'_0 is more plausible than S'_1 in both S'_0 and S'_1, i.e., $B(S'_0, 0, S'_0)$, $B(S'_1, 1, S'_0)$, $B(S'_0, 0, S'_1)$, $B(S'_1, 1, S'_1)$. Notice that S'_0 and S'_1 have the same belief structure, so the constraint violation mentioned above is resolved.

Unfortunately, under this new scheme we have a problem with beliefs about future beliefs. If we were to redefine *Bel* in the obvious way to accommodate the extra argument in B, our example would entail the very counterintuitive

[16] We can also handle a variant where nothing is believed about the cat sleeping initially by making the groups (1) and (2) the most plausible.

$Bel(\neg \text{LIGHT}_1 \wedge Bel(\text{LIGHT}_1, do(\text{SENSELIGHT}, Now)), S_0)$, i.e., in S_0, the agent believes that the light is not on but thinks that after sensing he will believe that it is on. Our approach—which uses a fixed plausibility ordering on situations and simply drops situations that conflict with sensing results from the B relation—avoids both of these problems.

Hunter [28, pp. 67–69] claims that revision operators $*$ as defined by Definition 32 are, strictly speaking, not functions since given two different situations t, t' where $K(t) = K(t')$ it is not always the case that $t * \phi = t' * \phi$. However, by Definition 32, $t * \phi$ is defined as $do(A_\phi, t)$ and so each revision function $*$ can be considered defined relative to a particular situation. We could, as it were, write this more precisely as $*_t$ however our interest here is to examine how closely our framework complies with the AGM framework and not to use it to define AGM-like revision operators.[17]

7. Comparison to other frameworks

One proposal that is related to ours is that of Demolombe and Pozos Parra [29]. Rather than reifying the accessibility relation in the style of Moore [11] and Scherl and Levesque [5,6] as we do here, they introduce belief modalities B_i and successor state axioms for these modal operators. For each modal operator B_i and fluent F two successor state axioms are required:

$$B_i\big(F\big(do(a,s)\big)\big) \equiv \Gamma^+_{i_1,F}(a,s) \vee B_i\big(F(s)\big) \wedge \neg\Gamma^-_{i_1,F}(a,s),$$

$$B_i\big(\neg F\big(do(a,s)\big)\big) \equiv \Gamma^+_{i_2,F}(a,s) \vee B_i\big(\neg F(s)\big) \wedge \neg\Gamma^-_{i_2,F}(a,s).$$

Furthermore, these modal operators are assumed to obey axioms for the modal logic KD. Each modal operator can be used to represent the beliefs of a different agent. In our framework this would be achieved through the introduction of accessibility relations B_i (as noted in [29]). One issue that Demolombe and Pozos Parra point to is that our approach assumes that every agent agrees on the same set of effects (i.e., successor state axioms) for each fluent. More precisely, our approach is directed towards representing and reasoning about the beliefs of a single agent. This agent can reason about actions performed by other agents as these are treated as exogenous actions. To deal with multiple agents, the axioms for each (precondition, successor state, etc.) would be grouped into separate action theories and reasoned about separately. This seems appropriate since each agent reasons about its own beliefs using its own conception of the world. This does not preclude that agents can also reason about other agents' beliefs. Demolombe and Pozos Parra [29] also consider that their approach might be better suited to dealing with noisy sensors.

In [30] Demolombe and Pozos Para propose another solution which adopts an accessibility relation along the lines of Moore [11] and Scherl and Levesque [5,6] however, in order to deal with multiple agents, they include a term for agents. More specifically, their accessibility relation is of the form $K(i, s', s)$ meaning that situation s' is compatible with agent i's beliefs in situation s. Their notion of belief $B(i, \phi(s', s), s', s)$ is defined as follows:

$$B\big(i, \phi\big(s', s\big), s', s\big) \overset{\text{def}}{=} \forall s'\big(K\big(i, s', s\big) \supset \phi\big(s', s\big)\big)$$

Their aim is to deal with belief change in the situation calculus without recourse to plausibilities. They introduce the notion of *real* and *imaginary* situations and the actions whose occurrence can be witnessed by an agent. Real situations correspond to the *actual* situations in which the agent is placed[18] while imaginary situations are simply those alternative situations which are compatible with the agent's beliefs. Each agent can have different successor state axioms for fluents and also for real and imaginary situations. The successor state axioms take the form

$$\begin{aligned}&\forall s\forall s'\forall a\forall \vec{x}\big(real(s) \supset \big(K\big(i, s', s\big) \supset \big(p\big(\vec{x}, do\big(a, s'\big)\big) \equiv\\ &\quad \Gamma^+_{i,p}\big(i, \vec{x}, a, s'\big) \vee \big(a = sense_p(i) \wedge p(\vec{x}, s)\big) \vee p\big(\vec{x}, s'\big) \wedge\\ &\quad\quad \neg\big(\Gamma^-_{i,p}\big(i, \vec{x}, a, s'\big) \vee \big(a = sense_p(i) \wedge \neg p(\vec{x}, s)\big)\big)\big)\big)\big)\end{aligned}$$

where $sense_p(i)$ is a sensing action informing i about the truth of $p(\vec{x}, s)$. The distinction between real and imaginary situations allows the agent to maintain K-related situations even when they do not accord with the agent's observations. Note however that as we have indicated in Section 4, the actual plausibility values themselves are not important and we provide a way of applying our framework without having to supply these values explicitly.

Another proposal for handling belief revision in the situation calculus is that of del Val and Shoham [31]. Their approach models beliefs through formulae of the form $holds(bel(\phi), s)$ and a causal axiom:

$$\forall s \forall \mu \big(\exists s'.holds\big(bel(\mu), s'\big)\big) \supset holds\big(bel(\mu), result\big(learn(\mu), s\big)\big).$$

They use a circumscription policy to reason about the effects of performing actions and to reason about beliefs. Both our approaches are characterized axiomatically, however theirs does contain an assumption that they did not axiomatize, namely

[17] Note that Darwiche and Pearl's [18] revision operators are defined similarly.

[18] As such, the successor of a real situation is also a real situation.

that every possible valuation of the fluents is witnessed by a situation. The advantages of our approach over theirs are that they do not handle belief introspection and our formal apparatus is much simpler than theirs.

Friedman and Halpern's [25] approach is to begin with a very general framework that combines state dynamics with belief, and to see what further constraints need to be placed in order to capture the standard approaches to belief revision and update. It is interesting to note that there are several points in common between our belief revision frameworks. We both generalized an existing framework for representing knowledge using possible worlds by adding a plausibility structure to the worlds. Both of our possible worlds contain the history of all actions or events until the current time. We both make the assumption that the set of accessible worlds at a given time is a subset of the accessible worlds at earlier times. Friedman and Halpern's framework, like ours, contains the assumption that the agent does not revise its assessment of the relative plausibility of situations. Rather, the agent is assumed to have a prior assessment of the relative plausibility of situations and the dynamics of the agent's beliefs arises from dropping possible worlds that conflict with new information. As a consequence, their agents also cannot recover from inconsistency. However, they consider this to be a problem with the postulates rather than with their framework.

Friedman and Halpern have a similar constraint to (TE). They require that accessible situations have the same plausibility structure. Both of our frameworks are synchronous in that accessible situations have the same "current time", and both our agents have perfect recall for past actions. One point of difference between our frameworks is that they need separate sets of constraints to obtain revision and update, so they cannot intersperse revisions and updates, whereas we can.

We have already noted above that the plausibility ordering remains fixed in our framework, yet this is sufficient to yield some rather desirable properties. However, several proposals for modifying plausibility orderings have been put forward in the literature (many stemming from the work of Spohn [20]). Boutilier [32] and Williams [33] propose a scheme whereby, upon receiving new information ϕ, the most plausible ϕ-worlds become the most plausible worlds while all other worlds retain their relative levels of plausibility. Spohn [20] (and Darwiche and Pearl [18]) adopt a method where ϕ and $\neg\phi$-worlds retain their relative levels of plausibility amongst themselves, but the two groups are "shifted" relative to each other, making the most plausible ϕ-worlds the most plausible worlds. However, these approaches do not consider belief introspection. While plausibilities themselves do not change in our framework, it must be kept in mind that the B-relation is also important in terms of determining belief and it may certainly change as a result of performing sensing actions. Furthermore, changes in belief can also be brought about by non-sensing actions which have the capacity to alter the environment.

The main aspect that distinguishes our work from previous approaches is the ability to represent belief introspection properties *within* the object language together with the facility to achieve iterated revision and update in a unified framework.

Belief change in the situation calculus has already been dealt with by Scherl and Levesque [6]. However, as noted previously, while they can handle belief update, they are limited to belief expansion. Del Val and Shoham [31] also address the issue of belief change in the situation calculus, and their theory deals with both revision and update. However, they cannot represent nested belief and consequently cannot deal with the issues of belief introspection and mistaken belief.

There are a variety of frameworks that accommodate both belief revision and belief update. As noted, this is one strength of the proposal by del Val and Shoham [31]. In a more traditional belief change setting, Boutilier [34] also provides a general framework that allows for both these forms of change. However, this framework cannot deal with introspection in the object language. One approach that supports both belief revision and update and also handles introspection is Friedman and Halpern [25]. Their approach to revision and update is fairly standard, but set within a very general modal logic framework that combines operators for knowledge, belief (interpreted using a plausibility ordering), and time. But they do not discuss interactions between revision and update and introspection. The work of Demolombe and Pozos Parra [29,30] can also handle belief introspection.

Another avenue of related work is that of modal logic accounts of belief change. Our account, taking some of its heritage from Moore [11], reifies the accessibility relation central to modal semantics. In an early work, Segerberg [35] developed a dynamic doxastic logic in which action modalities $[+\mathbf{B}\phi]$, $[-\mathbf{B}\phi]$ and $[*\mathbf{B}\phi]$ denote expansion, contraction and revision of the agent's belief state by $\mathbf{B}\phi$ respectively. This allows for formulas like $[*\mathbf{B}\phi]\mathbf{B}\psi$ with the meaning that revising the agent's belief state by belief in ϕ will result in belief in ψ. In Segerberg's framework, expansion, contraction, and revision are treated as actions working with belief formulas like $\mathbf{B}\phi$. This contrasts with our framework where there are only physical actions (leading to belief update) and sensing actions (leading to belief revision). More recent developments in this area include Herzig and Longin [36] who introduce a modal logic approach to this problem. Their language is based on propositional dynamic logic and introduces a modal belief operator *Bel* and the underlying logic is KD45. They introduce two successor state axioms:

$$\big(perc(a,b) \wedge \neg After_a \perp \wedge \neg BelAfter_b \perp\big) \supset (Feasible_a BelA \equiv BelAfter_b A),$$

$$\big(perc(a,b) \wedge \neg After_a \perp \wedge BelAfter_b \perp\big) \supset (Feasible_a BelA \equiv BelAfter_{enable_b} After_b A).$$

Their framework is capable of dealing with non-deterministic actions and misperception. Information is acquired when an *observation action observe*(ϕ) is performed. *observe*(ϕ) has ϕ as precondition. A "test action" *testIf*(ϕ) (similar to a sensing action) is treated as a non-deterministic choice between *observe*(ϕ) and *observe*$(\neg\phi)$. Thus in their framework, the need for belief revision arises when an action that is believed not to be executable is nonetheless perceived. As we can see in the second axiom above, this is handled by performing a special type of action $enable_a$ whose effect is to make action a

executable, i.e., make *a*'s preconditions true. This simple approach avoids the need for a plausibility ordering on epistemic alternatives. But only beliefs in the observed facts (and more generally the action's preconditions) are revised. These beliefs may have arisen due to other incorrect assumptions, but the latter would not be revised, unlike in a plausibility-based approach. It is not clear whether introspection about belief change is handled properly.

There has been a lot of work recently on this type of modal logic of knowledge/belief and action, for instance the work of van Ditmarsch et al. [37], where the paradigm of dynamic epistemic logic, a family of epistemic/doxastic logics with announcements and assignments/updates is developed. In van Ditmarsch et al. [38], an optimal regression method is developed for reasoning about knowledge and action within this type of propositional logic. However, most of this work does not deal with belief revision (one exception is Chapter 3 of [37], which is based on [35]). Van Benthem [39] also develops the correspondence between AGM-style belief change and dynamic epistemic logic. In particular, he shows how update rules can be used to modify the plausibility relation between possible worlds. Van Linder et al. [40] show how one can extend a propositional modal logic of multiagent knowledge, belief, and action to accommodate belief revision/contraction actions, following a similar approach to Segerberg's [35]. They also formalize agents' ability to perform actions (including having the required information) and how this applies to belief revision actions. However, they do not discuss how sensing actions might lead to belief revision. Thielscher [41] introduces a framework for knowledge in the *fluent calculus* [42]. He introduces a predicate $Knows(F, s)$ and provides knowledge update axioms to specify the effects of actions on the knowledge of the agent.

8. Conclusions and future work

We have proposed an account of iterated belief change that integrates into a well-developed theory of action in the situation calculus [4]. This has some advantages, in that previous work on the underlying theory can be exploited for dealing with issues such as solving the frame problem, performing automated reasoning about the effects of actions, specifying and reasoning about complex actions, etc. Our framework supports the introspection of beliefs and ensures that the agent is aware of when it was mistaken about its beliefs. Our account of iterated belief change differs from previous accounts in that, for us, the plausibility assignment to situations remains fixed over time. The dynamics of belief derives from the dynamics of the B modality and of the domain-dependent fluents. We showed that our theory satisfies all of the AGM, and the majority of the KM and DP postulates.

Our approach does have some limitations. In this paper, we have only looked at cases of belief change where the sensors are accurate, so that the agent only revises its beliefs by sentences that are actually true. It is the case that our successor state axiom for B ensures that the agent believes the output of its sensor after sensing. Also, our guarded sensed fluent axioms allow only hard (but context-dependent) constraints to be specified between the output of the sensor and the associated fluent; one cannot state that the sensor is only correct with a certain probability. However, we can also use beliefs to correlate sensor values to the associated fluents instead of guarded sensed fluent axioms. Thus, we could specify that the agent prefers histories where the sensors agree with the associated fluents more often to histories where they agree less often. Some of these issues are addressed in [26].

The fact that we never update the plausibility assignment, may suggest that our account has limited expressiveness. But we maintain that this is not the case. The example of Section 4 shows that we can handle some cases where a plausibility assignment update seems to be required.

We could extend the framework by having multiple agents that act independently and impart information to each other. Instead of beliefs changing only through sensing, they would also change as a result of *inform* actions. Shapiro et al. [43] provide a framework for belief expansion resulting from the occurrence of inform actions in the situation calculus, which we would like to generalize to handle belief revision.

Lakemeyer and Levesque [7] incorporate the logic of *only knowing* into the Scherl and Levesque framework of belief update and expansion. The traditional belief (and knowledge) operator specifies formulae that are believed (or known) by the agent, but there could be others. The 'only knows' operator is used to describe *all* that the agent knows, i.e., a formula that corresponds exactly to the knowledge state of the agent. In future work, we would like to define an analogous 'only believes' operator that could be used to describe exactly what the agent believes in a framework that supports belief revision as well as belief expansion.

Acknowledgements

The authors would like to gratefully acknowledge Craig Boutilier and Jim Delgrande for valuable comments on an earlier version of this paper. Financial support for the Canadian authors was gratefully received from the Natural Sciences and Engineering Research Council of Canada, and the Institute for Robotics and Intelligent Systems of Canada. The second author would like to acknowledge support from the Australian Research Council.

Appendix A. Proofs

Theorem 14.

$$\{\text{Axioms 1, 6, and 13}\} \models \forall s, s'.B(s', s) \supset (\forall s''.B(s'', s') \equiv B(s'', s)).$$

Proof. Suppose $M \models$ {Axioms 1, 6, and 13}. The proof is by induction on s. The base case follows directly from Axiom 13. Now suppose that for some situation S:

$$M \models \forall s'.B(s', S) \supset (\forall s''.B(s'', s') \equiv B(s'', S)). \tag{A.1}$$

We need to show that for any action A:

$$M \models \forall s'.B(s', do(A, S)) \supset (\forall s''.B(s'', s') \equiv B(s'', do(A, S))).$$

Let S''_1 be a situation such that $M \models B(S''_1, do(A, S))$. Then, by the successor state axiom for B (Axiom 1), there exists an S'_1 such that:

$$M \models B(S'_1, S) \wedge S''_1 = do(A, S'_1) \wedge (SF(A, S'_1) \equiv SF(A, S)). \tag{A.2}$$

We need to show that $M \models \forall s''.B(s'', S''_1) \equiv B(s'', do(A, S))$. We will prove the $\supset$ direction; the other case is similar. Suppose that for some situation S''_2, $M \models B(S''_2, S''_1)$. By the successor state axiom for B and (A.2), there exists a situation S'_2, such that:

$$M \models B(S'_2, S'_1) \wedge S''_2 = do(A, S'_2) \wedge (SF(A, S'_2) \equiv SF(A, S'_1)). \tag{A.3}$$

It follows from (A.1), (A.2), and (A.3) that $M \models B(S'_2, S)$. From this, (A.2) and (A.3), and the successor state axiom for B, it follows that $M \models B(S''_2, do(A, S))$. □

Lemma 20. *Let ϕ^* be domain-dependent formula, and A be a revision action for some formula ϕ. Then,*

$$\Sigma \models \forall s.\phi^*[s] \equiv \phi^*[do(A, s)].$$

Proof. By induction on ϕ^*. □

Theorem 21. *Let ϕ be a domain-dependent, uniform formula, and A be a revision action for ϕ wrt Σ. It follows that*:

$$\Sigma \models [\forall s.\phi[s] \supset Bel(\phi, do(A, s))] \wedge [\forall s.\neg\phi[s] \supset Bel(\neg\phi, do(A, s))].$$

Proof. We will prove the first conjunct; the proof of the second is similar. Let $M \models \Sigma$ and suppose that S and S'' are situations such that $M \models \phi[S] \wedge MPB(S'', do(A, S))$. By the successor state axiom for B, there is a situation S' such that $M \models B(S', S) \wedge S'' = do(A, S') \wedge (SF(A, S') \equiv SF(A, S))$. Therefore, since $M \models \phi[S]$ and A is a revision action for ϕ, $M \models \phi[S']$. By Lemma 20, $M \models \phi[S'']$. □

Theorem 22. *Let A be a revision action for a domain-dependent, uniform formula ϕ wrt Σ. The following set of sentences* (*which we denote by Γ*) *is satisfiable*:

$$\Sigma \cup \{Bel(\neg\phi, S_0), Bel(\phi, do(A, S_0)), \neg Bel(FALSE, do(A, S_0))\}.$$

Proof. Let S_1 and S_2 be situation constants. Since Σ does not contain initial state axioms, we can construct a model M of Σ such that:

$$M \models (\forall s'.B(s', S_0) \equiv (s' = S_1 \vee s' = S_2)) \wedge \neg\phi[S_1] \wedge \phi[S_2] \wedge pl(S_1) < pl(S_2).$$

It is easy to verify that $M \models \Gamma$. □

Theorem 24. *Let A be an update action for ϕ. Then*:

$$\Sigma \models \forall s.Bel(\phi, do(A, s)).$$

Proof. Let M be a model of Σ, and S, S'' be situations such that $M \models MPB(S'', do(A, S))$. By the successor state axiom for B, there is a situation S' such that $M \models S'' = do(A, S')$. Since A is an update action for ϕ, $M \models \phi[S'']$. □

Theorem 25. *Let A be a ground action term, and ϕ, ϕ' be uniform formulae. Then*:

$$\Sigma \models \forall s.Bel(\phi, s) \wedge \forall s' SF(A, s') \wedge (\forall s'.\phi[s'] \supset \phi'[do(A, s')]) \supset Bel(\phi', do(A, s)).$$

Proof. Let M be a model of Σ and S be a situation such that:

$$M \models Bel(\phi, S) \wedge \forall s' SF(A, s') \wedge (\forall s'.\phi[s'] \supset \phi'[do(A, s')]). \tag{A.4}$$

Suppose that for some situation S'':

$$M \models MPB\big(S'', do(A, S)\big). \tag{A.5}$$

We need to show that $M \models \phi'[S'']$. By the successor state axiom for B, there is a situation S' such that:

$$M \models B\big(S', S\big) \wedge S'' = do\big(A, S'\big). \tag{A.6}$$

Now, if we could show that $M \models MP(S', S)$, then the theorem would follow because we could infer from (A.4) that $M \models \phi[S']$ and also $M \models \phi'[S'']$. Suppose S'_1 is a situation such that:

$$M \models B\big(S'_1, S\big). \tag{A.7}$$

We need to show that $M \models pl(S') \leqslant pl(S'_1)$. It follows from the second conjunct of (A.4), (A.7), and the successor state axiom for B that $M \models B(do(A, S'_1), do(A, S))$. We can infer from this, (A.5), and (A.6) that $M \models pl(do(A, S')) \leqslant pl(do(A, S'_1))$. This and the successor state axiom for pl imply $M \models pl(S') \leqslant pl(S'_1)$, as desired. □

Theorem 26.

$$\Sigma \models \big[Bel(\phi, s) \supset Bel\big(Bel(\phi), s\big)\big] \wedge \big[\neg Bel(\phi, s) \supset Bel\big(\neg Bel(\phi), s\big)\big].$$

Proof. This follows directly from Theorem 14. □

Theorem 27. *Let A be a revision action for a domain-dependent, uniform formula ϕ wrt Σ. Then*:

$$\Sigma \models \forall s. Bel(\neg\phi, s) \wedge Bel\big(\phi, do(A, s)\big) \supset Bel\big(\mathbf{Prev}\big(\phi \wedge Bel(\neg\phi)\big), do(A, s)\big).$$

Proof. Let M be a model of Σ and S be a situation such that:

$$M \models Bel(\neg\phi, S) \wedge Bel\big(\phi, do(A, S)\big). \tag{A.8}$$

Suppose for some situation S'':

$$M \models MPB\big(S'', do(A, S)\big). \tag{A.9}$$

By the successor state axiom for B, there is a situation S' such that:

$$M \models B\big(S', S\big) \wedge S'' = do\big(A, S'\big). \tag{A.10}$$

We need to show that $M \models (\phi \wedge Bel(\neg\phi))[S']$. It follows from (A.8) and (A.9) that $M \models \phi[S'']$, and from this, (A.10), and Lemma 20 that $M \models \phi[S']$. Now, let S'_1 be a situation such that:

$$M \models MPB\big(S'_1, S'\big). \tag{A.11}$$

It follows from this, (A.10), and Theorem 14 that $M \models B(S'_1, S)$. If we could show that $M \models MP(S'_1, S)$ then the theorem would follow since we could infer from (A.8) that $M \models \neg\phi[S'_1]$. Let S'_2 be a situation such that $M \models B(S'_2, S)$. It follows from (A.10) and Theorem 14 that $M \models B(S'_2, S')$. We can now infer from (A.11) that $M \models pl(S'_1) \leqslant pl(S'_2)$. □

Theorem 34.

$$\Sigma \models \forall s B(s, s).$$

Proof. By induction on s. □

Theorem 36. *For any model M of Σ, the AGM postulates* (K*1)–(K*8) *are satisfied, when $*$ is defined as in Definition* 32 *for any situation t and domain-dependent uniform formulae ϕ and ψ.*

We prove this theorem by proving each postulate as a separate lemma.

Lemma 42 *(*K*1*). Under the assumptions of Theorem* 36, *$K(t * \phi)$ is deductively closed.*

Proof. This lemma follows from the fact that the *Bel* operator is closed over logical entailment. □

Lemma 43 *(*K*2*). Under the assumptions of Theorem* 36, *$\phi \in K(t * \phi)$.*

Proof. This follows directly from Theorem 21. $\square$

In the following lemma, we show conditions under which a most plausible, accessible situation remains so after a revision action for ϕ is performed.

Lemma 44. *Let A be a revision action for ϕ. Then, $\Sigma \models \forall s, s'.\phi[s] \wedge \phi[s'] \wedge MPB(s', s) \supset MPB(do(A, s'), do(A, s))$.*

Proof. Let $M \models \Sigma$ and t, t' be situations such that $M \models \phi[t] \wedge \phi[t'] \wedge MPB(t', t)$. Since A is a revision action for ϕ, $M \models SF(A, t) \wedge SF(A, t')$. Then, $M \models B(do(A, t'), do(A, t))$ follows from the successor state axiom for B. To see that $do(A, t')$ is most plausible, let t_1'' be a situation such that $M \models B(t_1'', do(A, t))$. By the successor state axiom for B, there is a situation t_1' such that $M \models t_1'' = do(A, t_1')$. By assumption, $M \models pl(t') \leqslant pl(t_1')$. By the successor state axiom for pl, $M \models pl(do(A, t')) \leqslant pl(t_1'')$. $\square$

Lemma 45 *(K*3). Under the assumptions of Theorem 36, $K(t * \phi) \subseteq t + \phi$.*

Proof. Suppose $\psi \in K(t * \phi)$, i.e., $M \models Bel(\psi, do(A, t))$. We need to show that $M \models Bel(\phi \supset \psi, t)$. Let t' be a situation such that $M \models MPB(t', t) \wedge \phi[t']$. Since $t * \phi$ is defined, $M \models \phi[t]$, therefore, it follows from Lemma 44 that $M \models MPB(do(A, t'), do(A, t))$. This, together with the hypothesis, imply that $M \models \psi[do(A, t')]$. It follows from Lemma 20 that $M \models \psi[t']$. $\square$

The following lemma identifies conditions under which the predecessor (under a revision action for ϕ) of a most plausible and accessible situation is also most plausible and accessible.

Lemma 46. *Let A be a revision action for ϕ. Then,*

$$\Sigma \models \forall s.\phi[s] \wedge \neg Bel(\neg\phi, s) \supset \bigl[\forall s''.MPB\bigl(s'', do(A, s)\bigr) \supset \exists s'.MPB\bigl(s', s\bigr) \wedge s'' = do\bigl(A, s'\bigr) \wedge \phi\bigl[s'\bigr]\bigr].$$

Proof. Suppose for some situation t, $M \models \Sigma \wedge \phi[t]$ and $M \not\models Bel(\neg\phi, t)$. Then, for some situation u, $M \models MPB(u, t) \wedge \phi[u]$. Further suppose that for some situation t'', $M \models MPB(t'', do(A, t))$. Since M satisfies the successor state axiom for B, there is a situation t' such that $M \models B(t', t) \wedge t'' = do(A, t') \wedge SF(A, t') \equiv SF(A, t)$. Since A is a revision action for ϕ, it follows that $M \models \phi[t']$. It remains to show that $M \models MP(t', t)$. Suppose to the contrary that there is a situation t^* such that $M \models B(t^*, t) \wedge pl(t^*) < pl(t')$. Since $M \models MPB(u, t)$, it follows that $M \models pl(u) \leqslant pl(t^*)$, and therefore $M \models pl(u) < pl(t')$. By the successor state axiom for pl, $M \models pl(do(A, u)) < pl(t'')$. It follows from the assumptions and Lemma 44 that $M \models B(do(A, u), do(A, t))$, which contradicts $M \models MPB(t'', do(A, t))$. $\square$

Lemma 47 *(K*4). Under the assumptions of Theorem 36, if $\neg\phi \notin K(t)$ then $t + \phi \subseteq K(t * \phi)$.*

Proof. Suppose $\neg\phi \notin K(t)$, i.e., $M \models \neg Bel(\neg\phi, t)$, and $\psi \in t + \phi$, i.e.,

$$M \models Bel(\phi \supset \psi, t). \tag{A.12}$$

We need to show that $M \models Bel(\psi, do(A, t))$. Suppose that for some situation t'', $M \models MPB(t'', do(A, t))$. Since $t * \phi$ is defined, $M \models \phi[t]$. Therefore, by Lemma 46, $M \models \exists s'.MPB(s', t) \wedge t'' = do(A, s') \wedge \phi[s']$. Let t' be a situation such that $M \models MPB(t', t) \wedge t'' = do(A, t') \wedge \phi[t']$. It follows from (A.12) that $M \models \psi[t']$. By Lemma 20, $M \models \psi[t'']$. $\square$

Lemma 48 *(K*5). Under the assumptions of Theorem 36, $K(t * \phi) \neq \mathcal{L}_{now}$.*

Proof. This follows directly from Corollary 35. $\square$

Lemma 49 *(K*6). Under the assumptions of Theorem 36, if $\models \phi \equiv \psi$, then $K(t * \phi) = K(t * \psi)$.*

Proof. This follows from the fact that the *Bel* operator preserves logical equivalence. $\square$

Lemma 50 *(K*7). Under the assumptions of Theorem 36, $K(t * (\phi \wedge \psi)) \subseteq (t * \phi) + \psi$.*

Proof. Suppose $\gamma \in K(t * (\phi \wedge \psi))$, i.e.,

$$M \models Bel\bigl(\gamma, do(A_{\phi\wedge\psi}, t)\bigr). \tag{A.13}$$

We need to show that $\gamma \in (t * \phi) + \psi$, i.e., $M \models Bel(\psi \supset \gamma, do(A_\phi, t))$. Suppose to the contrary that there is a situation t'' such that:

$$M \models MPB\big(t'', do(A_\phi, t)\big) \wedge (\psi \wedge \neg\gamma)\big[t''\big]. \tag{A.14}$$

Since $t * \phi$ is defined, $M \models \phi[t]$. Therefore, it follows from the successor state axiom for B and (A.14) that there is a situation t' such that:

$$M \models B\big(t', t\big) \wedge t'' = do\big(A_\phi, t'\big) \wedge \phi\big[t'\big]. \tag{A.15}$$

We can infer from Lemma 20, (A.14), and (A.15) that:

$$M \models (\psi \wedge \neg\gamma)\big[t'\big], \quad \text{and thus} \tag{A.16}$$

$$M \models \neg\gamma\big[do\big(A_{\phi\wedge\psi}, t'\big)\big]. \tag{A.17}$$

Now, it remains to show that $M \models MPB(do(A_{\phi\wedge\psi}, t'), do(A_{\phi\wedge\psi}, t))$, since this along with (A.17) contradicts (A.13). Since $t * (\phi \wedge \psi)$ is defined, $M \models (\phi \wedge \psi)[t]$. Therefore, the successor state axiom for B together with (A.15) and (A.16) imply that $M \models B(do(A_{\phi\wedge\psi}, t'), do(A_{\phi\wedge\psi}, t))$. Suppose t^{**} is such that:

$$M \models B\big(t^{**}, do(A_{\phi\wedge\psi}, t)\big). \tag{A.18}$$

We need to show that $M \models pl(do(A_{\phi\wedge\psi}, t')) \leqslant pl(t^{**})$. Since $M \models (\phi \wedge \psi)[t]$, it follows from the successor state axiom for B and (A.18) that there is a situation t^* such that:

$$M \models B\big(t^*, t\big) \wedge t^{**} = do\big(A_{\phi\wedge\psi}, t^*\big) \wedge (\phi \wedge \psi)\big[t^*\big]. \tag{A.19}$$

Similarly, it follows from the successor state axiom for B and (A.19) that $M \models B(do(A_\phi, t^*), do(A_\phi, t))$. From this and (A.14), we can infer that $M \models pl(t'') \leqslant pl(do(A_\phi, t^*))$. We can now use the successor state axiom for pl with (A.15) and (A.19) to infer that $M \models pl(do(A_{\phi\wedge\psi}, t')) \leqslant pl(t^{**})$ as required. □

Lemma 51 *(K*8). Under the assumptions of Theorem 36, if $\neg\psi \notin K(t * \phi)$, then $(t * \phi) + \psi \subseteq K(t * \phi \wedge \psi)$.*

Proof. Suppose that $\neg\psi \notin K(t * \phi)$, i.e.:

$$M \models \neg Bel\big(\neg\psi, do(A_\phi, t)\big), \quad \text{and} \tag{A.20}$$

for some formula γ, $\gamma \in (t * \phi) + \psi$, i.e.:

$$M \models Bel\big(\psi \supset \gamma, do(A_\phi, t)\big). \tag{A.21}$$

We need to show that $\gamma \in K(t * \phi \wedge \psi)$, i.e.:

$$M \models Bel\big(\gamma, do(A_{\phi\wedge\psi}, t)\big). \tag{A.22}$$

Suppose that for some situation t'':

$$M \models MPB\big(t'', do(A_{\phi\wedge\psi}, t)\big). \tag{A.23}$$

We need to show that $M \models \gamma[t'']$. Since $A_{\phi\wedge\psi}$ is a revision action for $\phi \wedge \psi$ and $t * (\phi \wedge \psi)$ is defined, it follows from the successor state axiom for B and (A.23) that there is a situation t' such that:

$$M \models B\big(t', t\big) \wedge t'' = do\big(A_{\phi\wedge\psi}, t'\big) \wedge (\phi \wedge \psi)\big[t'\big]. \tag{A.24}$$

We can infer from Lemma 20 that:

$$M \models (\phi \wedge \psi)\big[do\big(A_\phi, t'\big)\big]. \tag{A.25}$$

If we can show that $M \models MPB(do(A_\phi, t'), do(A_\phi, t))$, then by (A.21) and (A.25), $M \models \gamma[do(A_\phi, t')]$, and the theorem follows from Lemma 20 and (A.24), i.e., $M \models \gamma[t'] \wedge \gamma[t'']$. Since A_ϕ is a revision action for ϕ, it follows from (A.24) and the successor state axiom for B that $M \models B(do(A_\phi, t'), do(A_\phi, t))$. Now let t^{**} be a situation such that:

$$M \models B\big(t^{**}, do(A_\phi, t)\big). \tag{A.26}$$

It remains to show that $pl(do(A_\phi, t')) \leqslant pl(t^{**})$. From (A.20) and (A.26), it follows that there is a situation u'' such that:

$$M \models B\big(u'', do(A_\phi, t)\big) \wedge \psi\big[u''\big] \wedge pl\big(u''\big) \leqslant pl\big(t^{**}\big). \tag{A.27}$$

Since $t * \phi$ is defined, $M \models \phi[t]$. Therefore, since A_ϕ is a revision action for ϕ, it follows from (A.27), the successor state axiom for B, and Lemma 20 that there is a situation u' such that:

$$M \models B(u', t) \wedge u'' = do(A_\phi, u') \wedge (\phi \wedge \psi)[u']. \tag{A.28}$$

Since $t * (\phi \wedge \psi)$ is defined, $M \models \phi \wedge \psi[t]$. Therefore, since $A_{\phi \wedge \psi}$ is a revision action for $\phi \wedge \psi$, we can apply the successor state axiom for B again to yield:

$$M \models B(do(A_{\phi \wedge \psi}, u'), do(A_{\phi \wedge \psi}, t)).$$

This together with (A.23) implies that $M \models pl(t'') \leqslant pl(do(A_{\phi \wedge \psi}, u'))$. Using the successor state axiom for pl along with (A.24) and (A.28), we obtain $M \models pl(do(A_\phi, t')) \leqslant pl(u'')$. This, together with (A.27), yields $M \models pl(do(A_\phi, t')) \leqslant pl(t^{**})$, as desired. □

Lemma 39. *$w \in [K(t)]$ iff there exists $t' \in MPB(t)$ such that $w = Tr(t')$.*

Proof. The "if" part is obvious. For the "only if" part, we fix $w \in [K(t)]$. Since $\mathcal{L}_{now}$ is propositional and finite, let F be the conjunction of literals in w. Suppose, towards a contradiction, that there is no $t' \in MPB(t)$ such that $w = Tr(t')$. Then it is easy to see that $M \models Bel(\neg F, t)$. Therefore, $\neg F \in K(t)$, which implies $w \notin [K(t)]$. Contradiction. □

Theorem 40. *$K(t \diamond \phi)$ satisfies KM postulates* (K ◇ 1), (K ◇ 2), (K ◇ 4), (K ◇ 5), (K ◇ 8) *when ◇ is defined as in Definition* 37 *for any situation t and consistent domain-dependent uniform formulae ϕ and ψ.*

We prove this theorem by proving each postulate as a separate lemma.

Lemma 52 *(K ◇ 1). $K(t \diamond \phi)$ is closed.*

Proof. This lemma follows from the fact that the *Bel* operator is closed over logical entailment. □

Lemma 53 *(K ◇ 2). $\phi \in K(t \diamond \phi)$.*

Proof. This lemma follows directly from Theorem 24. □

After an update action is performed, the accessible situations are simply projected forward. In particular, as the following lemma shows, the most plausible accessible situations are preserved.

Lemma 54. *Let A be an update action for ϕ and t be a ground situation term. Then,*

$$\Sigma \models \forall s.MPB(s, t) \equiv MPB(do(A, s), do(A, t)).$$

Proof. This follows from the definition of an update action and the successor state axioms for B and pl. □

Lemma 55 *(K ◇ 4). $K(t \diamond \phi) = \mathcal{L}_{now}$ iff $K(t) \models FALSE$.*

Proof. This follows from Lemma 54. □

Lemma 56 *(K ◇ 5). If $\models \phi \equiv \psi$, then $K(t \diamond \phi) = K(t \diamond \psi)$.*

Proof. This follows from the definition of *ua*. □

Lemma 57 *(K ◇ 8). If $MPB(t) \neq \emptyset$ then $K(t \diamond \phi) = \bigcap_{t' \in MPB(t)} Tr(t' \diamond \phi)$.*

Proof. Let A denote $ua(\phi)$. Note that $K(t \diamond \phi) = \bigcap_{t' \in MPB(do(A,t))} Tr(t')$ and $t' \diamond \phi = do(A, t')$, therefore we need to show that:

$$\bigcap_{t'' \in MPB(do(A,t))} Tr(t'') = \bigcap_{t' \in MPB(t)} Tr(do(A, t')).$$

This follows from Lemma 54. □

Theorem 41. *Postulates* (DP1), (DP3) *and* (DP4) *are satisfied when $*$ is defined as in Definition* 32 *for any situation t and domain-dependent uniform formulae ϕ and ψ.*

We prove this theorem as a series of lemmas.

Lemma 58 *(DP1). Under the conditions of Theorem 41, if $\psi \models \phi$, then $K((t*\phi)*\psi) = K(t*\psi)$.*

Proof. Suppose $\gamma \in K((t*\phi)*\psi)$, i.e.:

$$M \models Bel\big(\gamma, do\big(A_\psi, do(A_\phi, t)\big)\big). \tag{A.29}$$

We need to show that $\gamma \in t*\psi$, i.e., $M \models Bel(\gamma, do(A_\psi, t))$. Let t'' be a situation such that

$$M \models MPB\big(t'', do(A_\psi, t)\big). \tag{A.30}$$

We need to show that $M \models \gamma[t'']$. Since A_ψ is a revision action for ψ and $t*\psi$ is defined, by the successor axiom for B, there is a situation t' such that

$$M \models B\big(t', t\big) \wedge t'' = do\big(A_\psi, t'\big) \wedge \psi\big[t'\big]. \tag{A.31}$$

Also, since $\psi \models \phi$, it follows that $M \models \phi[t']$. By Lemma 20 and (A.31), $M \models \psi[do(A_\phi, t')]$. Therefore, we can apply the successor state axiom for B twice to obtain $M \models B(do(A_\psi, do(A_\phi, t')), do(A_\psi, do(A_\phi, t)))$. Now, if it were also the case that $M \models MP(do(A_\psi, do(A_\phi, t')), do(A_\psi, do(A_\phi, t)))$, then the theorem would follow since we could infer from (A.29) that $M \models \gamma[do(A_\psi, do(A_\phi, t'))]$, and $M \models \gamma[t'']$ would follow from this, Lemma 20, and (A.31). Suppose to the contrary that there is a situation t^{***} such that:

$$M \models B\big(do\big(A_\psi, do\big(A_\phi, t'\big)\big), do\big(A_\psi, do(A_\phi, t)\big)\big) \wedge pl\big(t^{***}\big) < pl\big(do\big(A_\psi, do\big(A_\phi, t'\big)\big)\big). \tag{A.32}$$

Since A_ψ is a revision action for ψ, and $(t*\phi)*\psi$ is defined, it follows from the successor state axiom for B that there is a situation t^{**} such that:

$$M \models B\big(do\big(A_\phi, t^{**}\big), do(A_\phi, t)\big) \wedge t^{***} = do\big(A_\psi, t^{**}\big) \wedge \psi\big[t^{**}\big]. \tag{A.33}$$

We can infer from this and the successor state axiom for B that there is a situation t^* such that: $M \models B(t^*, t) \wedge t^{**} = do(A_\phi, t^*)$. It follows from this, (A.33), and Lemma 20 that $M \models \psi[t^*]$. Also, since $t*\psi$ is defined, $M \models \psi[t]$. Therefore, we can use the successor state axiom for B again to obtain: $M \models B(do(A_\psi, t^*), do(A_\psi, t))$. We can now see from (A.30) that $M \models pl(t'') \leqslant pl(do(A_\psi, t^*))$. However, it follows from (A.32), the situation equations, and repeated application of the successor state axiom for pl that: $M \models pl(do(A_\psi, t^*)) < pl(t'')$. Contradiction. $\square$

Lemma 59. *Under the conditions of Theorem 41, $\phi \in K((t*\phi)*\psi)$.*

Proof. We need to show that $M \models Bel(\phi, do(A_\psi, do(A_\phi, t)))$. Let t''' be a situation such that:

$$M \models MPB\big(t''', do\big(A_\psi, do(A_\phi, t)\big)\big). \tag{A.34}$$

Since A_ϕ is a revision action for ϕ and $t*\phi$ is defined, it follows from two applications of the successor state axiom for B that there is a situation t' such that:

$$M \models B\big(t', t\big) \wedge t''' = do\big(A_\psi, do\big(A_\phi, t'\big)\big) \wedge \phi\big[t'\big].$$

It follows from Lemma 20 that $M \models \phi[t''']$. $\square$

(DP3) follows as a corollary.

Corollary 60 *(DP3). Under the conditions of Theorem 41, if $\phi \in K(t*\psi)$, then $\phi \in K((t*\phi)*\psi)$.*

Lemma 61 *(DP4). Under the conditions of Theorem 41, if $\neg\phi \notin K(t*\psi)$, then $\neg\phi \notin K((t*\phi)*\psi)$.*

Proof. Suppose that there is a situation t^{**} such that:

$$M \models MPB\big(t^{**}, do(A_\psi, t)\big) \wedge \phi\big[t^{**}\big]. \tag{A.35}$$

We need to show that there is a situation t^+ such that:

$$M \models MPB\big(t^+, do\big(A_\psi, do(A_\phi, t)\big)\big) \wedge \phi\big[t^+\big].$$

Since A_ψ is a revision action for ψ and $t*\psi$ is defined, it follows from the successor state axiom for B, Lemma 20, and (A.35), there is a t^* such that:

$$M \models B\big(t^*, t\big) \wedge t^{**} = do\big(A_\psi, t^*\big) \wedge (\phi \wedge \psi)\big[t^*\big]. \tag{A.36}$$

Since $t * \phi$ and $(t * \phi) * \psi$ are defined, it follows from two applications of the successor state axiom for B that $M \models B(do(A_\psi, do(A_\phi, t^*)), do(A_\psi, do(A_\phi, t)))$. Now, if we can show that $M \models MP(do(A_\psi, do(A_\phi, t^*)), do(A_\psi, do(A_\phi, t)))$, then the theorem would follow because Lemma 20 implies that:

$$M \models \phi[t^*] \equiv \phi[do(A_\psi, do(A_\phi, t^*))],$$

and $M \models \phi[do(A_\psi, do(A_\phi, t^*))]$ follows from this and (A.36). Let t''' be a situation such that:

$$M \models B(t''', do(A_\psi, do(A_\phi, t))). \tag{A.37}$$

We need to show that $M \models pl(do(A_\psi, do(A_\phi, t^*))) \leqslant pl(t''')$. It follows from the successor state axiom for B, Lemma 20, and (A.37) that there is a situation t' such that:

$$M \models B(t', t) \wedge t''' = do(A_\psi, do(A_\phi, t')) \wedge \psi[t']. \tag{A.38}$$

We can use the successor state axiom for B again to yield:

$$M \models B(do(A_\psi, t'), do(A_\psi, t)).$$

From this and (A.35), it follows that $M \models pl(t^{**}) \leqslant pl(do(A_\psi, t'))$. Using (A.36), (A.38), and the successor state axiom for *pl*, we can infer that $M \models pl(do(A_\psi, do(A_\phi, t^*))) \leqslant t'''$, as desired. □

References

[1] S. Shapiro, M. Pagnucco, Y. Lespérance, H.J. Levesque, Iterated belief change in the situation calculus, in: A.G. Cohn, F. Giunchiglia, B. Selman (Eds.), Principles of Knowledge Representation and Reasoning: Proceedings of the Seventh International Conference (KR2000), Morgan Kaufmann Publishers, San Mateo, CA, 2000, pp. 527–538.

[2] J. McCarthy, Situations, Actions and Causal Laws, Stanford University Artificial Intelligence Project Memo 2, Stanford University, Stanford, CA, 1963.

[3] J. McCarthy, P.J. Hayes, Some philosophical problems from the standpoint of artificial intelligence, in: B. Meltzer, D. Michie (Eds.), Machine Intelligence, vol. 4, Edinburgh University Press, Edinburgh, 1969, pp. 463–502.

[4] R. Reiter, The frame problem in the situation calculus: A simple solution (sometimes) and a completeness result for goal regression, in: V. Lifschitz (Ed.), Artificial Intelligence and Mathematical Theory of Computation: Papers in Honor of John McCarthy, Academic Press, 1991, pp. 359–380.

[5] R.B. Scherl, H.J. Levesque, The frame problem and knowledge-producing actions, in: Proceedings of the Eleventh National Conference on Artificial Intelligence, AAAI Press/The MIT Press, 1993, pp. 689–695.

[6] R.B. Scherl, H.J. Levesque, Knowledge, action, and the frame problem, Artificial Intelligence 144 (1–2) (2003) 1–39.

[7] G. Lakemeyer, H.J. Levesque, AOL: A logic of acting, sensing, knowing, and only knowing, in: Principles of Knowledge Representation and Reasoning: Proceedings of the Sixth International Conference (KR-98), 1998, pp. 316–327.

[8] H.J. Levesque, What is planning in the presence of sensing?, in: Proceedings of the Thirteenth National Conference on Artificial Intelligence (AAAI-96), Portland, OR, 1996, pp. 1139–1146.

[9] P. Gärdenfors, Knowledge in Flux: Modeling the Dynamics of Epistemic States, MIT Press/Bradford Books, Cambridge, MA, 1988.

[10] H. Katsuno, A.O. Mendelzon, Propositional knowledge base revision and minimal change, Artificial Intelligence 52 (1991) 263–294.

[11] R.C. Moore, Semantical considerations on nonmonotonic logic, Artificial Intelligence 25 (1985) 75–94.

[12] G. De Giacomo, H.J. Levesque, Progression using regression and sensors, in: Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI-99), 1999, pp. 160–165.

[13] R. Reiter, Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems, MIT Press, Cambridge, MA, 2001.

[14] H.J. Levesque, F. Pirri, R. Reiter, Foundations for a calculus of situations, Electronic Transactions on Artificial Intelligence 2 (3–4) (1998) 159–178.

[15] F. Pirri, R. Reiter, Some contributions to the metatheory of the situation calculus, Journal of the ACM 46 (3) (1999) 325–361.

[16] C.E. Alchourrón, P. Gärdenfors, D. Makinson, On the logic of theory change: Partial meet contraction and revision functions, Journal of Symbolic Logic 50 (1985) 510–533.

[17] P. Peppas, A.C. Nayak, M. Pagnucco, N.Y. Foo, R. Kwok, M. Prokopenko, Revision vs. update: Taking a closer look, in: W. Wahlster (Ed.), Proceedings of the Twelfth European Conference on Artificial Intelligence (ECAI-96), 1996, pp. 95–99.

[18] A. Darwiche, J. Pearl, On the logic of iterated belief revision, Artificial Intelligence 89 (1–2) (1997) 1–29.

[19] C. Boutilier, Iterated revision and minimal change of conditional beliefs, Artificial Intelligence 25 (3) (1996) 262–305.

[20] W. Spohn, Ordinal conditional functions: A dynamic theory of epistemic states, in: W. Harper, B. Skyrms (Eds.), Causation in Decision, Belief Change, and Statistics, Kluwer Academic Publishers, 1988, pp. 105–134.

[21] D.K. Lewis, Counterfactuals, Harvard University Press, 1973.

[22] J. McCarthy, Epistemological problems in artificial intelligence, in: Proceedings of the Fifth International Joint Conference on Artificial Intelligence, Morgan Kaufmann Publishers, Los Altos, CA, 1977, pp. 1038–1044.

[23] J. McCarthy, Circumscription: A form of non-monotonic reasoning, Artificial Intelligence 13 (1–2) (1980) 27–39.

[24] J. McCarthy, Modality, si! modal logic, no!, Studia Logica 59 (1) (1997) 29–32.

[25] N. Friedman, J.Y. Halpern, A knowledge-based framework for belief change. Part II: Revision and update, in: J. Doyle, E. Sandewall, P. Torasso (Eds.), Proceedings of Knowledge Representation and Reasoning (KR-94), 1994, pp. 190–201.

[26] S. Shapiro, Belief change with noisy sensing and introspection, in: L. Morgenstern, M. Pagnucco (Eds.), Working Notes of the IJCAI-05 Workshop on Nonmonotonic Reasoning, Action, and Change (NRAC 2005), 2005, pp. 84–89.

[27] S. Shapiro, M. Pagnucco, Iterated belief change and exogenous actions in the situation calculus, in: R. López de Mántaras, L. Saitta (Eds.), Proceedings of the Sixteenth European Conference on Artificial Intelligence (ECAI-04), IOS Press, Amsterdam, 2004, pp. 878–882.

[28] A. Hunter, Belief change in the presence of actions and observations: A transition system approach, Ph.D. thesis, Simon Fraser University, Vancouver, BC, July 26, 2006.

[29] R. Demolombe, M.P. Pozos Parra, A simple and tractable extension of situation calculus to epistemic logic, in: Z.W. Ras, S. Ohsuga (Eds.), Proceedings of the Twelfth International Symposium on Intelligent Systems, in: LNAI, vol. 1932, 2000, pp. 515–524.

[30] R. Demolombe, M.P. Pozos Parra, Belief change in the situation calculus: A new proposal without plausibility levels, in: Proceedings of the Workshop on Belief Revision and Dynamic Logic at the Seventeenth European Summer School on Logic, Language and Information, 2005.

[31] A. del Val, Y. Shoham, A unified view of belief revision and update, Journal of Logic and Computation 4 (1994) 797–810.

[32] C. Boutilier, Revision sequences and nested conditionals, in: Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence (IJCAI-93), 1993, pp. 519–525.
[33] M.-A. Williams, Transmutations of knowledge systems, in: Proceedings of the Fourth International Conference on Principles of Knowledge Representation and Reasoning (KR-94), 1994, pp. 619–629.
[34] C. Boutilier, Generalized update: Belief change in dynamic settings, in: Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95), 1995, pp. 1550–1556.
[35] K. Segerberg, Belief revision from the point of view of doxastic logic, Bulletin of the IGPL 3 (1995) 535–553.
[36] A. Herzig, D. Longin, Sensing and revision in a modal logic of belief and action, in: F. can Harmelen (Ed.), Proceedings of the Fifteenth European Conference on Artificial Intelligence, IOS Press, 2002, pp. 307–311.
[37] H. van Ditmarsch, W. van der Hoek, B. Kooi, Dynamic Epistemic Logic, Synthese Library, vol. 337, Springer, 2007.
[38] H.P. van Ditmarsch, A. Herzig, T.D. Lima, Optimal regression for reasoning about knowledge and actions, in: Proc. of the Twenty-Second AAAI Conference on Artificial Intelligence, Vancouver, British Columbia, Canada, July 22–26, 2007, pp. 1070–1076.
[39] J.V. Benthem, Dynamic logic for belief revision, Journal of Applied Non-Classical Logics 17 (2) (2007) 129–155.
[40] B. van Linder, W. van der Hoek, J.-J.C. Meyer, Actions that make you change your mind, in: A. Laux, H. Wansing (Eds.), Knowledge and Belief in Philosophy and Artificial Intelligence, Akademie Verlag, Berlin, 1995, pp. 103–146.
[41] M. Thielscher, Representing the knowledge of a robot, in: A. Cohn, F. Giunchiglia, B. Selman (Eds.), Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth Conference, Morgan Kaufmann, 2000, pp. 109–120.
[42] M. Thielscher, Introduction to the fluent calculus, Electronic Transactions on Artificial Intelligence 2 (3–4) (1998) 179–192.
[43] S. Shapiro, Y. Lespérance, H.J. Levesque, Specifying communicative multi-agent systems, in: W. Wobcke, M. Pagnucco, C. Zhang (Eds.), Agents and Multi-Agent Systems—Formalisms, Methodologies, and Applications, in: LNAI, vol. 1441, Springer-Verlag, Berlin, 1998, pp. 1–14.